In compliance with the
Canadian Privacy Legislation
some supporting forms
may have been removed from
this dissertation.

While these forms may be included
in the document page count,
their removal does not represent
any loss of content from the dissertation.

University of Alberta

Detection and Diagnosis of Plant-Model Mismatch for Real-Time Optimization

by

Amar Bin Halim

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Master of Science**.

in

Process Control

Department of Chemical and Materials Engineering

Edmonton, Alberta

Fall 2003

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisisitons et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-87774-4
Our file Notre référence
ISBN: 0-612-87774-4

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

Canada

University of Alberta

Library Release Form

Name of Author: Amar Bin Halim

Title of Thesis: Detection and Diagnosis of Plant-Model Mismatch for Real-Time Optimization

Degree: Master of Science

Year this Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Date: JUNE 24, 2003

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Detection and Diagnosis of Plant-Model Mismatch for Real-Time Optimization** submitted by Amar Bin Halim in partial fulfillment of the requirements for the degree of **Master of Science** in *Process Control.*

Date: JUNE 24, 2003

Abstract

Automation has become a vital part of the modern process industry. Companies are striving to reduce production cost by efficient operation of their processes, while meeting environmental and safety regulations. Advanced process control strategies are playing a significant role in achieving that target. *Model-based on-line optimization*, also called *Real Time Optimization* (RTO), is a popular advanced control and optimization strategy. An RTO system uses a model to predict process performance, and uses that information to find and operate the process close to the optimum operating conditions. The success of an RTO system depends on the accuracy of the model, and any plant/model mismatch can cause loss of performance. This thesis proposes several practical methods for RTO system fault detection and diagnosis.

The proposed method uses *reduced profit gradient* information from the plant and the model to quantify and monitor the mismatch. The angle between the two profit gradients is used an an indicator of RTO performance and faults. The *error matrix* is calculated from the sensitivity matrices of the plant and the model. QR decomposition technique is used to investigate the orientation of the *error matrix* with respect to the *profit gradients*. A *profit weighted error matrix* is generated by scaling the *error matrix* with the profit gradients. Singular value decomposition is used to analyze this *profit weighted error matrix* to identify the source of plant-model mismatch.

The application of the proposed methods is demonstrated using simulation of a six heat-exchanger network model.

Acknowledgements

بسم الله الرحمن الرحيم

In the name of God, the Most Merciful, the Most Kind.

I begin by thanking God Almighty for His mercy and kindness. I am grateful to Him for making my parents the biggest blessing on my life.

I express my deepest gratitude to my supervisors, Professor J. Fraser Forbes and Professor P. James McLellan, for their extraordinary patience, support, and encouragement. It was a pleasure to build on their outstanding knowledge in the field of advanced control and optimization. A person can only dream of having two great people like them as supervisors.

Finally, I thank the Computer Process Control group of University of Alberta for their fruitful advice and support.

Contents

1 Introduction 1

1.1 RTO Overview . . . 3
1.2 Model-Based RTO . . . 4
1.3 Illustrative Example . . . 6
1.4 Problems and Issues in RTO . . . 9
1.5 Thesis Scope and Objective . . . 13
1.6 Thesis Conventions . . . 14
1.7 Assumptions . . . 15

2 Detecting RTO Performance Degradation 17

2.1 RTO System: An Optimization Problem . . . 18
2.2 Reduced Profit Gradient . . . 20
2.3 Plant-Model Mismatch & RTO Performance . . . 22
2.4 Root Causes of Plant-Model Mismatch . . . 24
2.5 Reduced Gradient Estimation Methods . . . 25
2.5.1 First-Order Perturbation-Based Method . . . 26
2.5.2 PRTS Based Method . . . 27
2.5.3 Time Series Model-Based method . . . 28
2.5.4 Modified PRTS Based Method . . . 29
2.6 Detection of Plant-Model mismatch . . . 31
2.7 Mismatch Detection Under Uncertainty . . . 34
2.8 Performance Evaluation Using Gradients . . . 38
2.9 Plant-Model Mismatch Cases . . . 40
2.10 Example Case Study . . . 43

3 Diagnosis of Plant-Model Mismatch 50

3.1 Tool Selection . . . 53
3.2 Diagnosis using Modified QR Decomposition . . . 58

3.2.1 Diagnostic Tools from Modified QR 59
3.2.2 Illustrative Example 62
3.3 Diagnosis using Singular Value Decomposition 67
3.3.1 Diagnostic Tools from SVD 71
3.3.2 Illustrative Example Revisited 75
3.4 An Integrated Diagnosis Approach 78
3.4.1 Steps Involved in Mismatch Diagnosis 78
3.4.2 Example 80
3.5 Summary 86

4 Case Study: Exchanger Network **88**
4.1 CASE STUDY I 90
4.1.1 Detection of Plant-Model Mismatch 91
4.1.2 Modified QR Decomposition 93
4.1.3 Singular Value Decomposition 95
4.2 CASE STUDY II 97
4.2.1 Modified QR Decomposition 99
4.2.2 Singular Value Decomposition 101

5 Summary, Future Work & Conclusions **104**
5.1 Summary 104
5.2 Future Work 107
5.3 Conclusion 109

Bibliography **110**

A Heat Exchanger Network Model **112**
A.1 Process Description 112
A.2 Parameter Estimation and Optimization 115
A.3 Process Equations 116
A.4 Simulation Results 118

B Case Study Model **122**
B.1 Process Description 122
B.2 Parameter Estimation and Optimization 124
B.3 Process Equations 125
B.4 Simulation Results 128

List of Figures

1.1 A Standard RTO loop 4
1.2 Illustrative heat exchanger example of RTO application 7
1.3 Surface temperature vs. heat transfer coefficient in a liquid heating system 8

2.1 RTO steps on a profit contour 23
2.2 Relation between mismatch angle and RTO performance 33
2.3 The confidence limit cones for plant and model 35
2.4 The largest and the smallest angles between the mismatch cones . . . 36
2.5 Information on the ellipses 37
2.6 Flow Chart 39
2.7 Model optimum between plant optimum and initial operating point . 41
2.8 Plant optimum between model optimum and initial operating point . 41
2.9 Model optimum between plant optimum and initial operating point . 42
2.10 Changes in the product of the gradients 43
2.11 Gradient vectors on profit contour from 4^{th}, 5^{th} & 6^{th} RTO steps . . . 44
2.12 Gradient vectors on profit contour from 7^{th}, 8^{th} & 9^{th} RTO steps . . . 45
2.13 Gradient vectors on profit contour from the 1^{st} to the 5^{th} RTO steps . 46
2.14 Gradient vectors on profit contour from the 6^{th} to the 9^{th} RTO steps 47

3.1 Influence of vector alignment with profit gradient $\frac{\partial P}{\partial y}$ on the overall profit 59
3.2 Illustrative Example of an Industrial Process 64
3.3 Graphical presentation of SVD decomposition of a matrix 70
3.4 Illustrative Example: Heat Exchanger Network 81

4.1 Heat Exchanger Network 89

A.1 Heat Exchanger Network (Numbers indicate stream number) 113
A.2 RTO setpoints and their corresponding value of the profit function . . 119
A.3 The two major measured variables (T_6 and T_7) 120

A.4 The plant-model mismatch angle at different RTO setpoints 120
A.5 Profit contour of the heat exchanger network 121
A.6 3D profit surface and contour of the heat exchanger network 121

B.1 RTO setpoints implemented by the optimizer 129
B.2 Value of the objective function at different RTO steps 130
B.3 Plant-model mismatch angle plotted against RTO steps 131

List of Tables

2.1 Comparison between different methods of reduced gradient estimation 31
2.2 RTO setpoints and their corresponding mismatch angles (Case I) . . 48
2.3 Operating point and their corresponding mismatch angle (Case II) . . 49

B.1 Flowrate and temperature of the streams 123

Chapter 1

Introduction

Competitive markets, higher energy costs and strict environmental regulations have pushed the process industries toward efficient production techniques. Advanced process control systems have assisted industry in achieving this target to a large extent. Better performance of control systems has given operators the confidence to operate plants closer to their physical constraints to increase productivity. The economic gain and operational reliability accompanying deployment of these systems has encouraged industry to invest further in advanced control systems. This drive has taken process control from single-loop controls to more complex multi-variable model predictive control systems and beyond.

Advanced control systems use models of the physical process to predict the process response to setpoint or manipulated-variable changes. To achieve further performance gains, economic information has been incorporated into these plant models in order to ensure that the control system considers the economic advantage of its control actions prior to their implementation. This technique of using a process model to determine the optimum operating policy is usually termed *real-time-optimization* (RTO), *on-line optimizing control*, or *off line process scheduling* (Forbes *et al.*, 1994). Until the 1980s, the lack of fast computer systems made it impossible to implement this kind of control strategy for large-scale systems (e.g., distillation columns, catalytic cracking units). With the advent of powerful computers and new control strategies, it is becoming more and more cost-effective to use these control schemes. In 1998 alone, approximately 250 online optimization systems were installed in the process

industries (White, 1998).

Over the years, keeping pace with the increasing computing capability, the complexity and size of process models has increased significantly. A typical model for a catalytic cracking unit with all its associated units can be as large as a hundred thousand equations (Darby and White, 1988). Debugging or upgrading these complex models is time consuming and expensive. Even so, the competitive advantage of these complex control strategies has encouraged industries to adopt them to maximize their profit margins. Adding to this already complex problem, the process industries regularly make changes in process feed, operating conditions, product quality, and so on to track changing market demand and take advantage of supply opportunities. Implementing these changes to complex control systems, like RTO systems, has become one of the main obstacles to smooth plant operation. Without an adequate maintenance effort, inaccurate estimation of process outputs from the model will direct the process towards less profitable operating conditions. In a large model, identifying the equations that need to be modified can be very tedious and time consuming.

There can be many factors that can change over the life of an RTO system. Small changes in the process can go unnoticed, and these changes create plant-model mismatch and thereby loss of RTO system performance. In the worst-case situations, the RTO system can make setpoint changes that reduce the generated-profit rather than increasing it. For these reasons, RTO users have recognized the need for techniques to:

- detect optimizer performance degradation,
- diagnose the reasons behind erroneous RTO moves.

Having tools available to perform these two tasks can make it easier and more reliable to operate an RTO system. This thesis provides procedures to assess the performance of the RTO system periodically over time. Furthermore, these procedures

can help determine the root cause of any performance loss whenever the RTO system fails to meet operating standards.

This thesis proposes numerical tools which use information available from the process model as well as from plant experiments to detect plant-model mismatch and to try to isolate the part of the model causing it. It is shown that comparing the reduced gradients of the plant and its model can be an efficient technique to ensure model adequacy within the current operation region, and to ensure acceptable operation of the overall optimization mechanism. In summary, the objective of this thesis is to investigate the impact of plant-model mismatch in RTO systems, and to propose novel techniques for detecting and diagnosing mismatch problems.

1.1 RTO Overview

RTO systems can be grossly categorized into two types (Forbes, 1994):

1. *Model-based methods*, which use a steady-state or a dynamic mathematical model to represent the process. The model is updated at regular intervals using process data to ensure accurate representation of the process and the objective function. Optimization algorithms are applied to the model to estimate the optimum operating point.

2. *Direct-search-based methods*, which use on-line plant experimentation to perturb the plant and directly measure the performance index. Process information is used to generate a performance-based profit-surface, which is used to determine the direction of the next RTO step. This technique may, in specific situations, give better predictions at the cost of longer plant-experimentation time. It is widely used in processes with short response times, e.g., electrical instruments such as missile guidance systems (Fu *et al.*, 1997)).

Of these two methods, model-based method is more popular in the process industry, as it requires fewer plant experiments. The time and cost related to plant

experimentation make direct search based methods less attractive for multivariate process applications.

1.2 Model-Based RTO

A simplified block diagram of a model-based RTO system is presented in Figure 1.1. The six basic blocks of the RTO loop (Forbes *et al.*, 1994) are elaborated below:

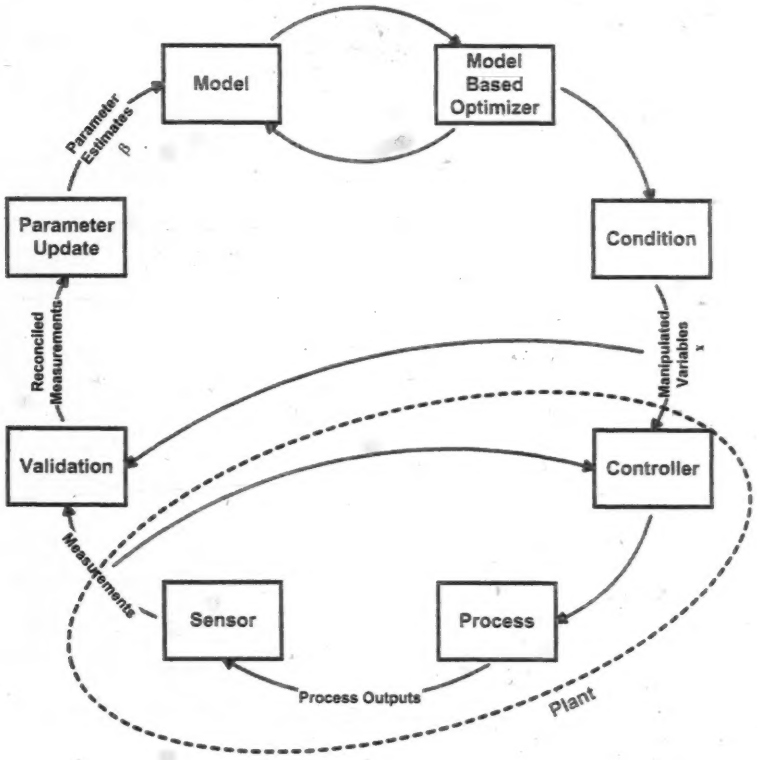

Figure 1.1: A Standard RTO loop

1. *Plant*: For the purpose of this thesis, the term *plant* describes the physical process as well as the controllers used to implement the *setpoints* determined

by the RTO system. Note that setpoints will be referred to as *operating points* interchangeably. The three basic components of the plant are:

(a) *The Process*: This is the physical process that is being controlled. It can be a complex process, like a crude-oil processing refinery, or a simple plug-flow reactor.

(b) *The Sensor System*: The sensors collect process measurements for control and optimization purposes.

(c) *The Control System*: These are the low-level controllers necessary for maintaining stable process operation. The setpoints calculated by the RTO system are implemented by these controllers.

2. *Data Validation Block*: Process and measurement noise are an integral part of the process measurements. Process measurements need to be filtered and *reconciled* to provide accurate estimation of the plant conditions, e.g., pressure, temperature, concentration (Darby and White, 1988).

3. *Process Model*: The model consists of a set of equations relating the process inputs and the outputs. There will also be constraint equations, limiting the maximum and the minimum value of these inputs and outputs. These constraints are a part of the model, and they must be satisfied to get satisfactory representation of the process performance. The model is used to estimate the process outputs for different operating conditions of the process. Using the input-output relations from the process model, the RTO system estimates the profitability of operating the process at any given operating condition.

4. *Parameter Update Block*: Processes are usually time-varying and nonlinear in nature. A fixed process model is unlikely to give satisfactory prediction of process performance beyond its designed operating conditions. An example of this would be a catalytic reactor. The activity of the catalyst changes with time, and these changes depend on the process feed rate, temperature and other operating conditions. As the activity of the catalyst changes, the model

has to be updated to match the process. This is usually done by modifying the crucial parameters, in this case catalyst activity, of the model using the available process operating data. An optimization routine may be used to estimate the model parameters that minimize the difference between model outputs and plant outputs for similar changes in their setpoints.

5. *Model-Based Optimizer*: The task of the optimizer is to predict the best operating conditions for the process, which increases profitability at every successive step. It uses an *objective function* or *profit function* to estimate the amount of profit generated at given setpoints. The optimizer finds the best set of operating conditions to maximize the objective function.

6. *Conditioning Block*: This block ensures that the setpoints proposed by the optimizer can be implemented without violating the physical constraints of the process.

1.3 Illustrative Example

A simple waste heat recovery system with two heat-exchangers is used to illustrate the difficulties associated with proper RTO system application (Forbes, 1994). The flow diagram of the system is presented in Figure 1.2.

The heat-exchange network shown in Figure 1.2 is used to recover heat from two hot streams (stream *11* and *13*) to heat a colder stream (stream *1*). The target is to recover the maximum amount of energy from the hot streams to economize the heating cost of the cold stream. The two-way-split valve is used to distribute stream *1* between the two exchangers. There is an unknown split-ratio that maximizes the amount of recovered heat from the hot streams.

It is evident that heat-transfer rates in the exchangers play a major role on the overall system. Factors, such as the flowrates of the hot and cold streams, heat-transfer areas, and fouling in the heat-transfer surfaces, can influence the overall

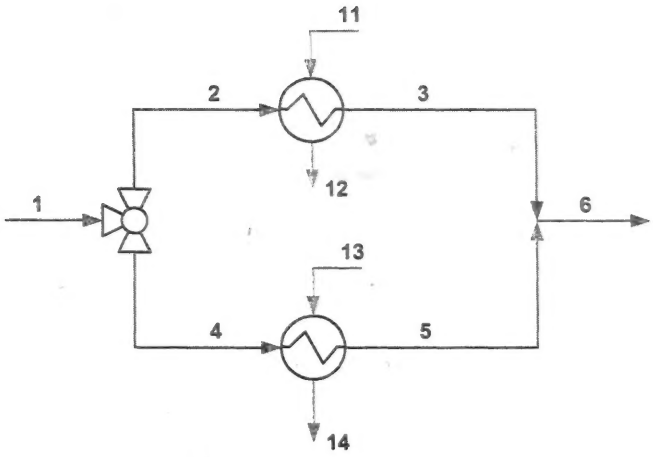

Figure 1.2: Illustrative heat exchanger example of RTO application

heat-transfer rates in the two heat-exchangers. Consequently, the flow-splitter valve, which changes the flow rate of the cold streams (stream 2 and stream 4), also influences the heat-transfer coefficients. Therefore, the process is highly coupled and may well be nonlinear depending on functional dependencies of overall-heat-transfer coefficient on the flowrates. If the RTO system is designed to update heat-transfer coefficients as opposed to having their functional dependence in the model, which are treated as constants for subsequent optimization calculations, then the RTO system may not be able to take the process to its true optimum operating condition (Forbes, 1994).

The situation can get even worse if the operating condition changes and the exchanger encounters a vapor phase in the incoming liquid streams. Lower heat conductivity of the vapor phase causes a reduction in the heat transfer coefficient (see Figure 1.3). If the model is not designed to consider vapor phase in the streams, the RTO system can make incorrect setpoint predictions.

It might be argued that this kind of problem can be resolved by using a better model to represents the process for both one-phase and two-phase flow conditions.

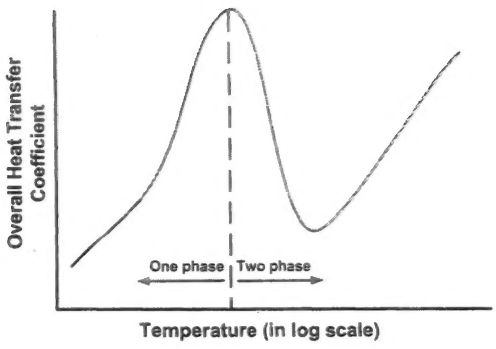

Figure 1.3: Surface temperature vs. heat transfer coefficient in a liquid heating system

Usually, RTO systems use simple models and may not incorporate phase changes until it is encountered.

The example clearly illustrates that a good model and a satisfactory parameter updating method cannot always ensure proper representation of a plant. It is necessary to have in-depth knowledge of the process mechanisms, which might not always be available. This is especially true when a process is operated under conditions that have not been encountered before. Therefore the model might fail to track the changing optima of the process with changing physical and environmental conditions. To solve this problem, the model parameters have to be updated to compensate for the plant-model mismatch. Parameter updating is conducted by rearranging the mathematical model to obtain equations relating the parameters with the measured variables and known thermodynamic, kinetic, and equilibrium properties (Cutler and Perry, 1983), and solving these equations using nonlinear regression, in where parameter values are adjusted to minimize squared prediction error (Bates and Watts, 1988).

Another example could be a distillation column. In its model it is assumed that all of the trays have liquid phase on them; however, sometimes due to high flow

rate of vapor from the bottom of the column the liquid can be carried over from the trays. This phenomena called, *vapor entrainment*, can cause trays to run dry. The dry trays do not work as a vapor-liquid equilibrium states, as considered in the model. Modelling this behavior in detail produces a numerically and computationally complex model, which is impractical for use in an RTO application. There is a trade-off between the model complexity and the ability to track process behavior in an efficient and timely fashion.

The examples discussed above illustrate that it is not possible to develop a "perfect model", which can represent a nonlinear process under all operating conditions. There will always be plant-model mismatch independent of complexity of the model and its parameter updating technique. The best solution is to use the best available model and to modify and update it to minimize profit loss.

1.4 Problems and Issues in RTO

The potential benefits of online optimization have been recognized by industry, as well as academia, for more than forty years. However, the implementation of these advanced control systems has seen limited success in most of the cases. Depending on the complexity and nonlinearity, the model of an industrial process is usually quite large and highly integrated. Successful application of an RTO system depends on a thorough understanding of the process and proper application of that knowledge in each of the components of the RTO system (Korchinski *et al.*, 1998). The size and complexity of the RTO system makes it very difficult to identify and diagnose faults in the system. Some of the reasons for RTO failure are elaborated below (White, 1998):

1. ***Online optimization benefits do not really exist***: In some cases, optimization systems have been implemented in processes where it is actually not needed. An example will be a time-invariant process. For such a process, it is sufficient to take the process to its optimum by trial and error or by offline computation and to continue operating under those conditions. In other cases,

changes in the plant or business environment can also render the optimization system unnecessary. Two cases where optimization is unnecessary are elaborated below:

- *Optimal solution does not change*: It is necessary for a process to have sufficient degrees of freedom for optimization. These degrees of freedom are usually control-variables which can be independently manipulated to change the operating conditions of the process. In some processes these major independent variables are set externally. In that case, there are insufficient degrees of freedom for the optimizer system and it cannot provide adequate performance. No change in RTO operating conditions with time can be an indication of such a situation.
- *Disturbance frequency too high*: The other problem can be having too much disturbance or process noise in the measured variables. In a steady-state model-based optimizer, steady-state process data is used to calculate the setpoints for the process. If large disturbances are occurring more rapidly than the plant settling time, it is seldom at steady state and the optimization system rarely executes. Under these conditions, online usage drops and the system is eventually turned off.

2. ***Inadequate optimization robustness***: Poor design can be a major contributor in the failure of an optimizer. Poorly designed systems exhibit lack of robustness and fail to function when operating conditions change. Three important reasons for lack of optimization robustness are elaborated below:

   - *No online-parameter updating*: Without parameter updating systems, process models cannot provide accurate predictions of process-response. Optimization projects are implemented using equipment-design data and plant-experiment data which are applied for initial parameter estimation. However, the plant response changes with time and those model parameters fail to represent the actual process behavior. Catalyst ageing, heat-exchanger

fouling, accumulation of inert substances in process loops, changes in feed source and composition, etc. can be major contributors to parameter mismatch. With time the prediction from the model and the actual plant response begins to diverge. Successful online optimization systems incorporate online model auto-calibration procedures that use actual plant data to regularly update the model parameters, and thereby model responses stay congruent with the plant. Systems have to incorporate this feature to operate for longer duration of time.

- *No compensation for process instrumentation problem*: In a real process, all sensor readings are not always reliable. Without regular calibration, they fail to give accurate readings. Further more, manipulated variables can become saturated, or an operator might override the basic control system. The optimization system must be able to recognize these problems and be able to function in a degraded mode when they happen. If the default option for these problems is simply not to execute the optimizer estimated setpoints, then the setpoints can rarely be executed. In such a case, the users might soon tire of relying on the optimizer and the system might fall into disuse.

- *Implementation on an unreliable infrastructure*: Optimization systems are like a chain of equipment and software, where the weakness of any one of the segments can cause failure to the overall system. The system must extract data from the process measurement system, perform its calculations and return the answers to the process. Users, both plant operators and system maintenance staff, must be able to determine what the system is doing and make changes, if necessary, through the provided user interfaces. Problems in any one of the parts will disable the optimization system. This can include network problems, user interface problems, computer system or database problems, and so forth.

3. ***Lack of unit feed information*** When designing an industrial process, engi-

neers try to minimize the number of measuring instruments to save money. This reduces the capital investment of the process at the cost of process controllability or observability. The lack of redundancy makes it difficult to reconcile the data and thereby minimize the effect of process disturbances. Due to the high cost of maintenance and operation of flow meters, a large process train might have only one or two flow measurements. Most of the intermediate flows in a process have to be approximated from knowledge of the process. These estimates show large deviation from the actual conditions when operating conditions change, and thus contribute to significant RTO system performance degradation over time.

4. ***Organizational disagreements on prices used***: In an industrial environment, like a refinery, the price of the feed (crude oil) and the products (e.g., gasoline, diesel) are not always exactly known. The process units share the same utilities and consumption of utilities by each of the units is rarely measured. Furthermore, recycle streams carry low-grade products from the end of the process to units at the front. These variations in pricing of the process streams and the interaction between them make it difficult to quantify the exact amount of value added by each unit.

5. ***Maintenance, support not budgeted with enough resources***: An RTO system works on the assumption that all the sensors are giving accurate readings within their limit of accuracy or that sufficient sensors are available to accurately detect gross errors and reconcile the measurement data, and all the controllers are well-tuned. Due to lack of proper maintenance and support, these criteria might not be met. Failure at the lower level of the control hierarchy definitely affects the performance of the advanced control system sitting at the top.

1.5 Thesis Scope and Objective

The performance criterion for a successful RTO system should be its ability to unambiguously identify and quickly converge to the optimum operating conditions for the plant. This description of RTO system performance requirement can be divided into three key points (Zhang *et al.*, 2002):

- minimize offset from optimal plant operation,
- reduce transmission of uncertainties around the RTO loop, and
- fast convergence of the RTO system

Even though these system performance goals are widely known, sufficient effort has not been made to utilize them in RTO performance improvement. Extended design cost (Zhang *et al.*, 2002) has been proposed to choose the best model for an RTO system. It is a technique that can be used only at the development stage of an RTO system. Once the RTO system is operational, it is impossible to change the complete model without shutting down the system. It is necessary to have analytical tools to detect and diagnose performance degradation in an operational RTO system. These tools should help the control engineers to conduct a performance check of the RTO system with minimal amount of disruption to the regular process operation. In addition, there should be tools to help engineers identify the source of the performance loss using offline calculation techniques. In this thesis, some novel ideas have been proposed to develop the tools necessary:

1. to detect when model fidelity of an RTO system has decreased to unacceptable levels;
2. to isolate the parts of the model that are contributing to that loss of performance.

1.6 Thesis Conventions

This section documents the conventions adopted in this thesis to simplify discussion. In this thesis, **plant** refers not only to the physical process but also to its control systems that enforce the setpoints determined by the RTO system. It is assumed that the plant is *stable* and can be operated in any point within some required operating range. The system of equations describing the material and energy balances, thermodynamic relationships, and so forth, along with the process operating constraints, is defined as the **model**.

The process variables are divided in to two categories: (a) **Manipulated variables (x)**, or the independent process variables, that can be independently manipulated for optimizing plant performance, and (b) **Dependent variables (y)**, whose values become uniquely defined once the values of the manipulated variables are set.

Plant sensitivity matrix refers to the sensitivity of the plant output to the plant inputs. Similarly, **model sensitivity matrix** is the first derivative of the model outputs with respect to the model inputs. Using the concept of a reduced space (Fletcher, 1987), the total number of plant inputs and outputs will be reduced to only the independent inputs. The sensitivity matrix of the plant and the model in the reduced space are referred as **reduced plant sensitivity matrix** and **reduced model sensitivity matrix** respectively. In this thesis, *plant sensitivity matrix* and *reduced plant sensitivity matrix*, and *model sensitivity matrix* and *reduced model sensitivity matrix* are used interchangeably.

At any operating point, **plant profit** refers to the actual profit that would be generated by operating the plant at those operating conditions. **Profit gradient** refers to the amount of change in profit for an infinitesimal change in each of the inputs and the outputs.

Throughout this thesis, all terms are explained on first usage.

1.7 Assumptions

The main assumptions made in this thesis are elaborated below. Any other assumptions related to this thesis are explained where used. The tools and concepts developed in this thesis use profitability as the performance evaluation criterion. The concepts are applicable for a process already in operation along with its RTO system. The objective of this thesis is to analyze the performance of the RTO system, and to diagnose the root cause of performance degradation.

The primary assumptions for this thesis are:

1. The process has a global optimum, which can be reached by standard optimization routines in a finite number of RTO steps.

2. The best possible control technique has been adopted beneath the optimization layer to ensure stability of the process and to enforce fast convergence to the new operating conditions.

3. The controlled process is *stable* and any setpoints proposed by the RTO system can be implemented on the process without causing an upset condition.

4. All the process measurements are bias-free and have Gaussian white noise added to them. There are no gross errors in the process data at the optimization level.

5. Economic information related to the process streams (e.g., purchase price of raw materials and utilities, selling price of products, value added in each processing step) are exactly known.

6. The plant gradients can be estimated from available plant data, and plant experiments can be conducted to minimize uncertainty when needed. The model gradients can be calculated by solving the equations in the model along with the constraints.

7. The profit gradient with respect to the process inputs and outputs ($\partial P/\partial \mathbf{x}$ and $\partial P/\partial \mathbf{y}$) should be the same for each of the plant and model.

8. The process parameters are considered to remain constant and independent of the manipulated variables, while evaluating the process gradients.

The following two chapters present techniques for plant-model mismatch detection and diagnosis. Chapter 2 introduces a gradient based tool for monitoring the performance of an RTO system. In Chapter 3, *QR decomposition* and *singular value decomposition* (SVD) are used to analyze the mismatch between the sensitivity matrices of the process and its model to identify the part of the model contributing to the loss of process performance. In Chapter 4, the proposed techniques are applied on a heat-exchanger network to evaluate their capabilities as well as their short-comings. Chapter 5 gives a brief summary of the overall thesis and proposes some ideas for further expansion of this work.

Chapter 2

Detecting RTO Performance Degradation

Real Time Optimization (RTO) is an on-line optimization technique used to increase the total productivity of a process by predicting its best operating conditions. *Model-based RTO* is the most popular form of industrial RTO systems, and uses a model of the process to estimate the direction of increasing profitability (Pfaff, 2001). The precision of finding the right direction depends on, among other factors, the accuracy of the model in predicting the process outputs for given operating conditions. Therefore, an inadequate model can cause the RTO system to drift away from the optimal operating conditions, and reduce the productivity of the process (Forbes *et al.*, 1994). Although RTO is a well-established process optimization technique, there is no procedure to ensure accuracy of the setpoints predicted by the RTO system. The lack of performance evaluation tools has contributed to the demise of many RTO systems.

The optimization algorithm of an RTO system generates sensitivity information relating the process inputs and outputs by a *sensitivity matrix*. The current form of the RTO system does not use this information explicitly in setpoint calculation (see **Optimization** and **Command Conditioning** blocks in Chapter 1). This sensitivity matrix provides useful information on the sensitivity of the model outputs to changes in the input variable setpoints. Similar sensitivity information can be obtained for the plant through plant experimentation. The sensitivity matrix of a good process model

will be similar to that of the process itself. In other words, any discrepancy between the plant and its model will create mismatch between the sensitivity matrices. This thesis uses the sensitivity matrix as a tool for plant-model mismatch detection and isolation, and proposes several mathematical techniques for analyzing the mismatch based on its effect on the total profitability of the plant.

This chapter proposes a method to compare the gradients from the plant and the model using a new concept of *plant-model mismatch angle*. It will be shown that the plant-model mismatch angle can be a simple and effective off-line tool for mismatch detection, which can improve the reliability of the RTO system.

2.1 RTO System: An Optimization Problem

At the core of an RTO system is the optimization problem, where process operating conditions are determined that maximize productivity, as well as meet the operating constraints and product specifications. For example, an RTO system for a *heat exchanger network* (see Chapter 5) will try to maximize heat recovery, thereby minimizing the heating and cooling cost of the streams. At the same time, it will ensure that the process streams are not heated or cooled beyond their operating constraints.

The general form of the optimization problem for the RTO system is written as,

$$\begin{aligned} &\underset{\mathbf{x},\mathbf{y}}{\text{minimize}} \; P(\mathbf{x},\mathbf{y},\boldsymbol{\beta}) \\ &\text{subject to:} \\ &\quad \mathbf{f}(\mathbf{x},\mathbf{y},\boldsymbol{\beta}) = 0 \\ &\quad \mathbf{g}(\mathbf{x},\mathbf{y},\boldsymbol{\beta}) \geqslant 0 \end{aligned} \tag{2.1}$$

where:

$\mathbf{x}$ = independent variables

$\mathbf{y}$ = dependent variables

$\boldsymbol{\beta}$ = model parameters

$P(\mathbf{x},\mathbf{y},\boldsymbol{\beta})$ = objective or profit function

$\mathbf{f}(\mathbf{x},\mathbf{y},\boldsymbol{\beta})$ = equality constraints

$\mathbf{g}(\mathbf{x},\mathbf{y},\boldsymbol{\beta})$ = inequality constraints

The independent variables ($\mathbf{x}$) are the manipulated variables, which can be changed independent of each other, to optimize the process. The dependent variables ($\mathbf{y}$) are the process measurements, which are related to the independent variables ($\mathbf{x}$) by the model equations or the process itself. For a process, usually the number of dependent variables ($\mathbf{y}$) will be larger than the number of manipulated variables ($\mathbf{x}$). Fewer independent variables ($\mathbf{x}$) means less degrees of freedom and reduced flexibility in process optimization. The examples studied in this thesis (see chapter 2, 3 and 4) have at least two degrees of freedom for process optimization.

The equality and inequality constraints (i.e. $\mathbf{f}(\mathbf{x},\mathbf{y},\boldsymbol{\beta}) = 0$ and $\mathbf{g}(\mathbf{x},\mathbf{y},\boldsymbol{\beta}) \geqslant 0$ respectively) constitute the model of the process. These equations are used to predict the process measurements ($\mathbf{y}$) for different values of the manipulated variables ($\mathbf{x}$). For *time-invariant* processes, these equations remain the same for any operating time. However, most processes are *time-varying* in nature, and the physical parameters of the process change with time and operating conditions. Examples of time-varying process parameters can be the fouling-factor of a heat-exchanger and the reactivity of a catalyst. The model parameters have to be updated at regular intervals by some external mechanism (see **Model Updater** in Chapter 1) for acceptable estimation of the process behavior using the process model. These parameters are referred to as the *model parameters*, ($\boldsymbol{\beta}$). The model parameters should be similar to the process parameters for satisfactory estimation of the process outputs.

The profit function, $P(\mathbf{x},\mathbf{y},\beta)$, relates the profitability of the plant with its input-output variables. For graphical visualization, the profit function can be presented as a profit surface in an $n+1$ dimension space (where n is defined as the total number of independent variables). The extra co-ordinate is the profit function ($P(\mathbf{x},\mathbf{y},\beta)$) evaluated at the corresponding values of the independent variables. This plot, is

referred to as a *profit surface* (Beveridge and Schechter, 1970). Figure 2.1 presents the profit-surface of a process with two manipulated variables. The profit-surface has local extrema as well as a global extremum (e.g. maxima or minima or both). The objective for the RTO system is to move the process operating conditions closer to the maximum on the profit surface.

2.2 Reduced Profit Gradient

As explained in the previous section, the RTO system is used to determine the operating conditions that can increase overall process profitability in successive RTO steps. It uses the process model in conjunction with the profit function to estimate the operating condition which can generate more profit than the current setpoints. The efficiency of the RTO system in finding the process optimum depends on the accuracy of all the blocks in the RTO loop (see Chapter 1). Each of the blocks in the loop depends on the blocks preceding it, and any fault in one of these blocks can contribute to inaccurate prediction of the process optimum. Therefore, the failure of any of the blocks will degrade the performance of the RTO system.

Referring to Equation 2.1, it can be noted that the profit function ($P(\mathbf{x}, \mathbf{y}, \boldsymbol{\beta})$) is a function of the manipulated ($\mathbf{x}$) and the measured variables ($\mathbf{y}$) of the process. This thesis assumes that the price factors (e.g., the price of all the raw materials, products, utilities) related to each of the inputs and outputs are exactly known (see **Assumptions** in Chapter 1). Therefore, the unknown quantities in the profit function are the input-output variables of the process. Considering the profit function to be continuous for all values of the input-output variables, the exact differential (Protter and Morrey, 1964) of the profit function ($P(\mathbf{x}, \mathbf{y}, \boldsymbol{\beta})$) is,

$$dP = \frac{\partial P}{\partial \mathbf{x}} d\mathbf{x} + \frac{\partial P}{\partial \mathbf{y}} d\mathbf{y} \tag{2.2}$$

Equation 2.2 assumes that $\boldsymbol{\beta}$ remains constant for small perturbations of the ma-

nipulated variables ($\mathbf{x}$). The sensitivities of the profit function to changes in the manipulated and measured variables are given by $\frac{\partial P}{\partial \mathbf{x}}$ and $\frac{\partial P}{\partial \mathbf{y}}$ respectively. Rearranging Equation 2.2 gives the total derivative of the profit function with respect to the manipulated variables ($\mathbf{x}$):

$$\frac{dP}{d\mathbf{x}} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \frac{d\mathbf{y}}{d\mathbf{x}} \tag{2.3}$$

Since $\frac{\partial P}{\partial \mathbf{x}}$ and $\frac{\partial P}{\partial \mathbf{y}}$ are considered to be known, the only unknown in this Equation 2.3 is $\frac{d\mathbf{y}}{d\mathbf{x}}$, the sensitivity of the process measurements to perturbation in the manipulated variables. This matrix ($\frac{d\mathbf{y}}{d\mathbf{x}}$) relating the process input-output variables is defined as the *gain matrix* or *sensitivity matrix*. Estimating the sensitivity matrix ($\frac{d\mathbf{y}}{d\mathbf{x}}$) is the most important and difficult element in any process control and optimization scheme. For most nonlinear processes, the gains change with changing operating conditions and they have to be re-calculated in every RTO step for efficient RTO operation.

To simplify the optimization problem, the equality constraints ($f(\mathbf{x}, \mathbf{y}, \boldsymbol{\beta}) = 0$) relating the independent variables are used to reduce the total number of variables available to the optimizer. This reduces the dimension of the optimization problem, and thereby the dimension of the profit gradient vector (∇P). This new profit gradient with lesser dimension is defined as the *reduced profit gradient*, $\nabla_r P$, where the subscript 'r' is used to indicate the reduced dimension of the problem. The concept of reduced gradient has been adopted from Fletcher and Gill (Fletcher, 1987; Gill *et al.*, 1981).

Based on Equation 2.3, for a plant with independent manipulated variables $\mathbf{x}$, plant measurements $\mathbf{y}$, and the profit or objective function $\mathbf{P}(\mathbf{x}, \mathbf{y})$; the reduced gradient of the objective function can be expressed as,

$$\nabla_r P|_{plant} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} \tag{2.4}$$

Similarly, at the same operating conditions the reduced gradient from the model will be,

$$\nabla_r P|_{model} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model} \tag{2.5}$$

Since the profit gradients with respect to the process inputs and outputs are exactly known (see **Assumptions** in Chapter 1), $\frac{dP}{d\mathbf{y}}$ and $\frac{dP}{d\mathbf{x}}$ will be the same for both the plant and the model. The only unknown terms in the Equations 2.4 and 2.5 are the sensitivity matrices of the plant and the model (i.e. $\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant}$ and $\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model}$). The plant sensitivity matrix is estimated by perturbing the process, while the model sensitivity matrix is calculated from the model equations. For a perfect model of a plant the two matrices are identical, subsequently any mismatch between the two can contribute to mismatch in the reduced profit gradients and is the sole source of mismatch.

2.3 Plant-Model Mismatch & RTO Performance

As pointed out in the previous section, the reduced profit gradient shows the sensitivity of the profit function to the decision variables. For efficient RTO operation, it is necessary that the profit gradient from the model be "similar" to the actual plant gradient. This is very unlikely in an actual process implementation and this can be traced back to mismatch between the gain matrices of the plant and its model. Therefore, when the reduced plant gradient is not equal to the reduced model gradient, within the limits of statistical significance, it will be said that the RTO system has a *plant-model mismatch.*

An illustrative example is used to explain the effect of plant-model mismatch on the performance of an RTO system. Consider the process profit-surface in Figure 2.1 with two manipulated variables ($\mathbf{x}_1$ and $\mathbf{x}_2$). The profit generated for any given setpoint of the manipulated variables ($\mathbf{x}_1$ and $\mathbf{x}_2$) is presented along the vertical axis. In most cases the optimal operating condition is not known. The objective for the

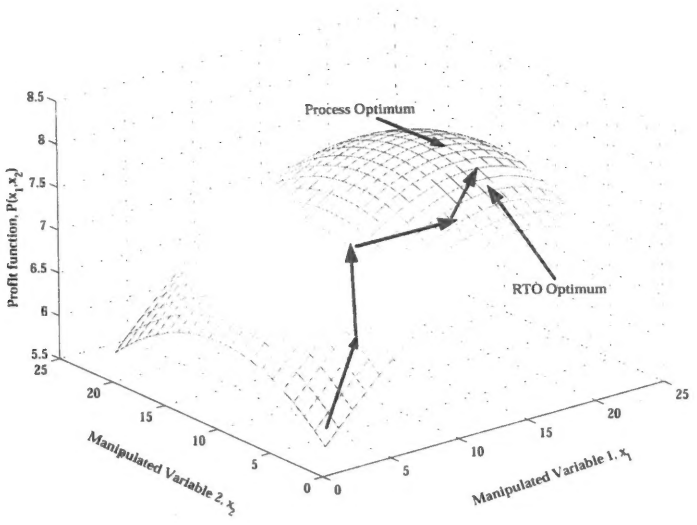

Figure 2.1: RTO steps on a profit contour

optimizer is to gradually move the process closer to this point in successive RTO steps. If the model-based optimization system could accurately estimate the profit, it could easily move the process to this optimum operating point. Since the profit is not known *a priori*, the optimizer uses the process model to estimate the shape of the surface and location of the optimal operation. As explained earlier, no model can be an exact representation of a process and there will always be some plant-model mismatch. Therefore, at any operating condition, the profit for the model will be different from the plant profit surface. The extent of the difference between the two surfaces will depend on the accuracy of the model in predicting the plant output variables ($\mathbf{y}$) for changes in the inputs ($\mathbf{x}$). Since the RTO system uses the model to predict its setpoint changes, mismatch in the profit surface will direct the process away from its true optimum operating point. Figure 2.1 shows the profit surface for the plant. The profit surface of the model changes at every RTO step, and so it is not shown. The arrows over the profit surface represent the RTO steps taken by the optimizer. In every RTO step the optimizer is estimating a direction of profit gain

and driving the process in that direction. In the first two RTO steps, the RTO system is successful in finding a direction of improving profit. In the third and fourth RTO steps, the RTO system starts to drift away from the true optimum. As expected, the optimizer converges to the model optimum, instead of the true optimum of the plant. The difference between the two points (the model and process optima) might be small, but when this small profit loss is accumulated over the entire run time of an RTO system it becomes quite significant. It is necessary to take measures to minimize this profit loss by eliminating the plant-model mismatch.

It is evident from Figure 2.1 that in each of the RTO steps the optimizer is trying to generate more profit than its previous operating condition. This characteristic of the optimizer ensures that the process is moving in an uphill direction over the profit surface, and guarantees that the RTO system is working properly. The RTO system fails when the up hill direction predicted by the RTO system is a downhill move for the actual process. In such a situation, comparing the gradient directions (estimated from the plant and the model) will show the correctness of the prediction made by the RTO system.

2.4 Root Causes of Plant-Model Mismatch

It is evident from the illustrative example given in section 1.3 that the performance of the RTO system depends heavily on model accuracy. In presence of a plant-model mismatch, the RTO system may move the process away from the optimal operating conditions. Therefore, it is necessary to understand the root causes of such plant-model mismatch. Plant-model mismatches can be grossly classified into two categories:

- **Parametric mismatch**: This type of mismatch is observed when the model structure is correct but the values of some of the plant parameters (e.g. catalyst activity, fouling factor, feed composition, and so forth) are incorrect.

- **Structural mismatch**: Structural mismatch is the case where model equations do not accurately represent all the process phenomena for any parameter value. In such case, the model will perform satisfactorily for some operating conditions, and will fail for some others. An example of structural mismatch would be the use of a linear model for a nonlinear process. The linear model may perform very well near the point of linearization, but it will fail as the process moves away from that point.

In general, industrial process models suffer from both of these types of mismatch. No model can accurately represent a non-linear process under all operating conditions. It might be necessary to modify the model to adapt to changes in the process. Furthermore, the physical parameters of the process (e.g., activity of the catalyst, fouling of the exchangers) will change over time and the model updater has to accurately estimate them for acceptable performance of the model.

2.5 Reduced Gradient Estimation Methods

Comparing the reduced gradient of the plant with the reduced gradient of the model is an effective method of plant-model mismatch detection. It is necessary to devise methods to estimate them with sufficient accuracy for efficient detection of plant-model mismatch. The reduced gradient of the model can be easily obtained from the model equations, and this can be done offline. On the other hand, estimating the reduced gradient of the plant can be quite difficult. Very little work has been done in reduced gradient estimation for the RTO system (Mansour and Ellis, 2003). There are, however, a number of techniques that can be modified to extract reduced gradient information from plant experiment data. This includes work by Roberts, Bamberger and Isermann, McFarlane and Bacon, and Golden and Ydstie (Roberts, 1979; Bamberger and Isermann, 1977; McFarlane and Bacon, 1989; Golden and Ydstie, 1989).

2.5.1 First-Order Perturbation-Based Method

A first-order perturbation method for parameter estimation was proposed by Roberts (Roberts, 1979). This algorithm is considered to be the foundation of the family of ISOPE (Integrated System Optimization and Parameter Estimation) algorithms. In the paper by Mansour and Ellis the authors state that, "*The real process derivatives $\frac{\partial y}{\partial x}$ will have to be obtained by perturbing the inputs, measuring the corresponding changes in output and applying finite-difference formulae.*" (Mansour and Ellis, 2003). To implement this technique, step changes are applied to each of the decision variables one at at time, and their effect on the steady-state output values are monitored. The input-output data is used to calculate the gradient relating the process outputs with the inputs. After implementing each of the step changes, the process is brought back to its previous state before applying another step change.

Advantage:

- Perturbing the manipulated variables ($\mathbf{x}$) one at a time eliminates the possibility of one manipulated variable influencing the estimate of the derivative with respect to another variable, which is quite prominent when multiple variables are perturbed simultaneously.

- The process is brought back to the original steady-state after each perturbation. This ensures that the accuracy of the gradients are not hampered by the transient behavior (i.e., inverse response, oscillation) of the process.

Disadvantage:

- For processes with many manipulated variables ($\mathbf{x}$), estimating the process gradients ($\frac{d\mathbf{y}}{d\mathbf{x}}$), one element at a time for each input is very time consuming. The process has to reach a steady state before and after a perturbation is applied. For a process with three inputs, it has to reach six steady states to get the full gain matrix of the process. In a time-varying process with many inputs, the actual process gradient might have changed during the estimation procedure. This method may be too slow to be applied in a large industrial process.

- The method is primarily applicable for unconstrained processes. Existence of constraints can complicate its application.
- Second-order derivative information ($\frac{d^2y}{dx^2}$) of the process input-output variables cannot be estimated from this method. Second-order derivative or, the Hessian of the variables can help understand the curvature of the *profit surface*.
- Information gathered during previous RTO intervals is ignored.
- Conventional experiment design emphasizes the use of joint moves to estimates sensitivity information. Perturbing the manipulated variables "one at a time" can make this method less effective for model parameter estimation.

2.5.2 PRTS Based Method

The method for process gradient estimation proposed by Bamberger and Isermann (1977) employs both first-order and second-order process derivative information to achieve faster convergence to the process optimum. In their proposed method, *Pseudo Random Ternary Signals* (PRTS) are added to the process inputs. Unlike *Random Binary Signals* (RBS), PRTS have three levels or distinct steady states in the process output, which helps gather second-order derivative information.

Once the process achieves steady operating conditions, a PRTS is added to all the process inputs simultaneously. The process data are fitted to a nonlinear dynamic model to estimate the static gain and second-order gradient information. The optimizer uses such gradient information to determine the next optimum operating point.

Advantage:

- Use of second-order information in Newton-based methods ensures faster convergence towards the optimum.

- Simultaneous perturbation of all the inputs greatly reduces the required time to perform plant experiments.

Disadvantage:

- The number of parameters in the process model (*Hammerstein* model) is much larger than other methods. More process data is required to calculate these parameters with acceptable statistical accuracy.
- This method uses the *instrumental variable* (IV) method for estimating the parameters of the Hammerstein model. The IV method may fail to give unbiased estimates in presence of non-stationary noise, and this can lead to inaccurate gradient estimation (Mansour and Ellis, 2003).
- As the process approaches its optimum, smaller RTO steps are taken. These smaller RTO steps result in a smaller *signal-to noise-ratio* (S/N) (Seborg *et al.*, 1989). When this noisy data is used for parameters estimation, the accuracy of the parameters may be unsatisfactory. This is particularly true in the estimation of second-order information since its accuracy is strongly influenced by the process noise. RTO approaches using second-order information (i.e., Newton-based methods), will make the process "jump around" its true optimum instead of staying at that operating point (Mansour and Ellis, 2003).

2.5.3 Time Series Model-Based method

McFarlane and Bacon (1989) presented two strategies of online gradient estimation. Similar to the methods proposed by Bamberger and Isermann, both of these methods add *pseudo random binary signal* (PRBS) to the process input signals to identify the model. The input-output data is fitted to a linear *autoregressive exogenous inputs* (ARX) model to approximate the steady state gains.

In these methods the perturbation signals are added throughout the operation of the process and do not wait for the process to reach steady-state operating conditions. This provides more process data for better parameter estimation. However, the estimated steady-state gradients are less reliable since they are estimated from transient operating data of the process.

Advantage:

- The perturbation signals are added to the inputs continuously, and this generates more process data for better estimation of the process gradients.
- The RTO steps are implemented by ramping the process from its previous setpoints to the new setpoints, and the perturbation signals are added to these ramp inputs. Since the process is in a transient state when the perturbation signals are applied, the calculated gradients will correspond to some interim operating condition between the two setpoints. This is similar to taking some weighted average of the gradients at the two operating conditions (i.e., the initial operating condition and the new operating condition).

Disadvantage:

- The effect of transient behavior of the process gets incorporated in the ARX model. This can contribute to biased estimation of the process gradient in presence of unusual process behaviors (e.g. inverse response, oscillation) if is does not properly account it.

2.5.4 Modified PRTS Based Method

The method proposed by Golden and Ydstie (1989) can be considered as an extension of Bamberger and Isermann's technique of gradient estimation. This method assumes having *a priori* knowledge of the static nonlinear behavior of the process. A nonlinear model is used to fit the process over a wider range of operating conditions than what

is possible with a linear or quadratic Hammerstein model. Similar to the previous method, PRTS signals are added to the process inputs as perturbation signals, while the process is ramped from one operating condition to the next. The accuracy of the gradients for the new operating point will depend on how long the process is maintained at the new steady state and also on the "forgetting factor" used in the algorithm.

A second-order Hammerstein model is fitted to the process data using a recursive least squares technique. The accuracy of the gradients depend on the amount of process data available for model fitting. This technique is relatively more computationally intensive.

Advantage:

- This technique provides better estimation of the process gradients.

Disadvantage:

- Fitting a larger non-linear model with more parameters to the process data can be computationally expensive.
- A large volume of process data required to get reliable estimate of the non-linear model parameters.

Each of the methods elaborated here have their advantages and shortcomings. Selection of the method for any application will depend on the process itself. In this thesis, Robert's method was used for the heat exchanger network problems. This method is relatively slow, but for the heat-exchanger networks, it gives better estimates of the gradients than other methods. Since the heat exchanger networks have fast response time, so large number of plant experiments can be conducted in a short time span.

Table 2.1: Comparison between different methods of reduced gradient estimation

	First Order Perturbation Based Method	**PRTS Based Method**	**Time Series Model Based Method**	**Modified PRTS Based Method**
Accuracy of the estimated gradients	Very accurate	Accurate	Accurate	Less accurate
Time required to estimate gradient	Very long	Fast	Fast	Moderate
Second-order information	Not estimated	Estimated	Not estimated	Estimated
Model used in estimation	Linear	Non-linear	Linear	Non-linear

2.6 Detection of Plant-Model mismatch

It is evident from the discussion in Section 1.3 that for most RTO systems plant-model mismatch is unavoidable. The only available solution to this problem is to keep the mismatch within acceptable limits. The first step toward that target is to develop a tool for detecting plant-model mismatch. This section proposes a technique for plant-model mismatch detection, which can help to distinguish between profitable and non-profitable RTO moves.

For an RTO system the primary objective is to generate more profit at every RTO step than the previous operating condition. As long as this condition is met, the RTO will climb up the profit surface toward the plant optimum. Since the actual curvature of the plant profit surface is not known to the RTO system, it has to rely on the process model to predict the process outputs and thereby the rate of change in profit with respect to setpoint changes. The accuracy of an RTO move depends on the precision of the model in predicting the gradient of profit at that operating point.

Referring back to Equation 2.4 and 2.5, the reduced profit gradients represent the sensitivity of the profit function to setpoint changes. The profit gradient is tangent

to the profit contour (Figure 2.1), showing the direction of maximum profit change. Hence, any RTO move in the direction of the profit gradient of the plant ($\nabla_r P|_{plant}$) will generate incrementally more profit, and take the process closer to its optimum. Ensuring that the reduced gradient of the model ($\nabla_r P|_{model}$) is colinear with that of the plant, may ensure a correct RTO move. Considering the two gradients as vectors, the RTO system is expected to perform better when the angle between the gradient vectors is small. Therefore, examining the angle between the two gradient can be useful for RTO performance evaluation.

The technique used in this thesis is to consider the gradients as vectors in an n dimensional space (where $n = dim(\mathbf{x})$) with a common origin and to determine the angle between the vectors. Using the definition of an inner product of two vectors, the following relation for the two profit gradients holds:

$$\left\langle \nabla_r P|_{plant}, \nabla_r P|_{model} \right\rangle = \left\| \nabla_r P|_{plant} \right\|_2 \| \nabla_r P|_{model} \|_2 \cos(\theta) \tag{2.6}$$

Rearranging Equation 2.6 for θ will give the angle between the two reduced profit gradients (see Equation 2.7). This angle will be referred as the *plant-model mismatch angle*, θ.

$$\theta = \cos^{-1}\left(\frac{\left\langle \nabla_r P|_{plant}, \nabla_r P|_{model} \right\rangle}{\left\| \nabla_r P|_{plant} \right\|_2 \| \nabla_r P|_{model} \|_2} \right) \tag{2.7}$$

As mentioned in Section 1.2, when the profit function is available, the only unknown term necessary to estimate the profit gradients are the reduced sensitivity matrices of the plant and the model (i.e., $\frac{dy}{dx}|_{plant}$ and $\frac{dy}{dx}|_{model}$). Once the reduced sensitivity matrices of the process and its model are calculated, the control engineer can conduct this simple calculation (Equation 2.7) to analyze the performance of the RTO system.

Besides plant-model mismatch angle, this thesis will also use the *inner product of the reduced profit gradients* to quantify the scale of the mismatch (see Equation 2.8).

The gradients are normalized to remove scale effects between different cases. The normalized inner product of the profit gradients, Φ can be written as follows:

$$\Phi = \left(\frac{\nabla_r P|_{plant} \cdot \nabla_r P|'_{model}}{\left\| \nabla_r P|_{plant} \right\|_2 \| \nabla_r P|_{model} \|_2} \right) \tag{2.8}$$

For a perfect model of a process, the reduced profit gradient of the model is colinear with the gradient of the plant, and the mismatch angle is zero. On the other hand, in the worst case situation the gradient point exactly the opposite direction, and for that case the mismatch angle is 180°. Depending on the accuracy of the model in representing the plant, the mismatch angle can vary between 0° and 180°. As the mismatch angle changes from 0° and 180°, the RTO system performance changes from good to critical.

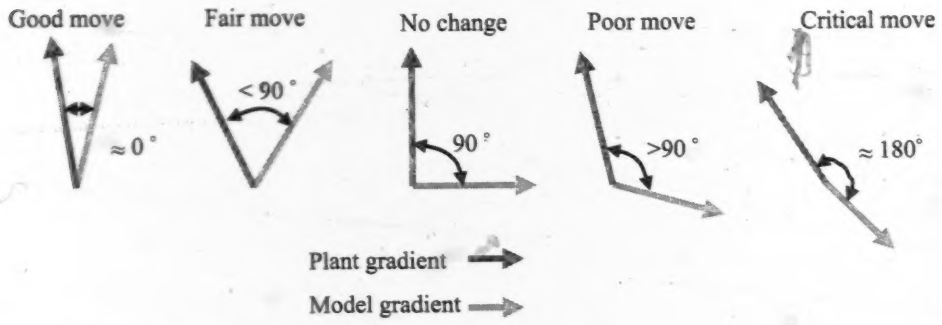

Figure 2.2: Relation between mismatch angle and RTO performance

A mismatch angle of 90° can be considered as the switch-over point. It indicates that the optimizer is moving the plant orthogonal to its plant profit gradient, and for a differentially-small RTO move the plant profitability remains constant. A mismatch angle slightly less than 90° indicates that the RTO may be marginally generating profit, but there is scope for significant performance improvement. Similarly, an

angle greater than 90° indicates that the RTO is slowly driving the process away from its optimum.

The plant-model mismatch angle can be a simple but effective tool for RTO performance evaluation. This technique reduces the burden of studying the different fragments of information (i.e., reduced sensitivity matrix, profit function) to draw any conclusion about the "health" of the RTO system. Its simplicity makes it easer to communicate to the plant operator, who can use it to understand the performance of the RTO system without going into the mathematical details of the system.

As mentioned earlier, the other method of plant-model detection is to look at the inner product of the profit gradients (Φ), which present the same information as the mismatch angle, but in a different format. Depending on the value of Φ, the following conclusions can be drawn:

- positive = generally moving in the right direction (mismatch angle $< 90°$)
- zero = may be moving in the direction of no change in profit (mismatch angle $= 90°$) or, one of the two gradients is zero, i.e., the current operating point is the plant or the model optimum.
- negative = process is moving in the wrong direction (mismatch angle $> 90°$)

The inner product of the gradients in not self-explanatory like the mismatch angle, but it is more convenient when used for any kind of numerical analysis.

2.7 Mismatch Detection Under Uncertainty

In the previous sections it is assumed that the reduced gradients can be estimated without any uncertainty in the process measurements. In practice, all the measurements are corrupted by process noise. The model parameters ($\boldsymbol{\beta}$) estimated from the

noisy process data are not exact, and have confidence bands associated with them. The sources of this noise can include external disturbances, raw material quality changes and sensor noise. Using expensive and sophisticated sensors and data reconciliation techniques will reduce some of the high frequency noise, but will not be able to completely eliminate it. For simplicity, this thesis will assume all process noise to be Gaussian. Based on this assumption, the probability density function of the profit gradient can be written as,

$$f_{(\nabla_r P)} = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{(\nabla_r P - \overline{\nabla_r P})' \Sigma^{-1} (\nabla_r P - \overline{\nabla_r P})}{2}} \tag{2.9}$$

where, $\overline{\nabla_r P}$ and Σ are the mean and the covariance matrix of the profit gradient $(\nabla_r P)$. The probability density function of ∇P is represented as $N(\overline{\nabla_r P}, \Sigma)$. Thus, the joint confidence region for the gradient estimates can be written as:

$$\left(\nabla P - \overline{\nabla P}\right)' \Sigma^{-1} \left(\nabla P - \overline{\nabla P}\right) = c^2 \tag{2.10}$$

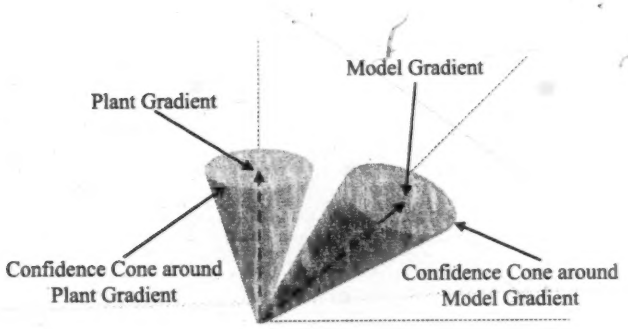

Figure 2.3: The confidence limit cones for plant and model

The implication of the above assumption is that, the profit gradient can be any line connecting the origin with a point on or in the hyper-ellipse described by Equation

2.10. Connecting the origin with the outer periphery of the ellipsoid generates a cone around the mean profit gradient vector as shown in Figure 2.3.

Figure 2.3 presents two cones, one for the plant and the other for the model profit gradient, with their means ($\overline{\nabla_r P}$) shown at the center of the cones. The size of these cones will depend on the accuracy of the profit gradient estimation techniques. For better estimation of the profit gradients, the cross-section of the cones will have smaller volume. In general, the influence of the process noise is more dominant in the plant than the model, and therefore the cone around the plant profit gradient should usually have larger diameter.

With uncertainty cones around the profit gradient vectors, it is not sufficient to consider the angle between the means of the gradients (i.e., $\overline{\nabla_r P}|_{plant}$ and $\overline{\nabla_r P}|_{model}$). It is necessary to investigate the best and worst case situation of the plant-model mismatch (see Figure 2.4). The best case is the smallest angle between the two cones. Similarly, the worst case is the largest angle subtended by the two cones. If the largest angle is less than 90°, then statistically it can be stated that the possible mismatch angles will be less than 90° at the level of probability used to calculate c^2 in Equation 2.10.

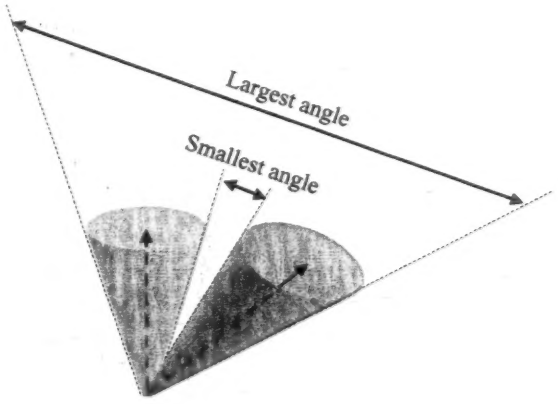

Figure 2.4: The largest and the smallest angles between the mismatch cones

It is necessary to find the two points on the two ellipses that constitute the largest mismatch angle. These two points will be at the two extremes of the ellipses and the distance between them will be larger than the distance between any other pair of points on the ellipses. The distance between any two points on the two ellipses can be described by a quadratic equation with multiple solutions. Solving the equation for the largest distance can be quite difficult. As an alternative, the search region for these two extremum points can be reduced using the geometric definition of the ellipses.

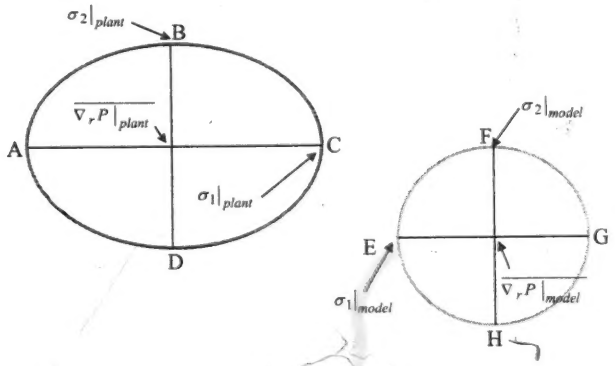

Figure 2.5: Information on the ellipses

Figure 2.5 shows the two confidence ellipses for the plant and the model. The center of the two ellipses are the mean profit gradients of the plant and the model ($\overline{\nabla_r P|_{plant}}$ and $\overline{\nabla_r P|_{model}}$). For a diagonal variance-covariance matrix (Σ), the length of the axes are given by the square root of the diagonal elements ($\sqrt{\Sigma_{ii}}$). The co-ordinates of the points where the axes intersect the ellipses (i.e., A, B, C, D, E, F, G, and H in Figure 2.5), can be calculated from the co-ordinate of the center and the length of the axes. The distances between these points (e.g., AE, AF, AG, AH, BE, BF, and so on) are calculated and the pair with the largest distance is chosen to be the initial pair of points for finding the two extremum points on the ellipses. The total number of distances to be compared is quite small (i.e., for a problem of dimension

$n = dim(\mathbf{x})$, the total number of distances is 2^n). Finding this ideal starting point for the search ensure fast convergence to the extremum points on the ellipses.

As discussed before, the inner product of the profit gradients (Φ) is close to one when the mismatch angle is small, and it is close to negative one for maximum possible mismatch. Using the equation of the ellipses as equality constraints, minimization of the product of the profit gradients helps determine the maximum angle between the two cones:

$$\underset{\mathbf{x}}{\text{minimize}} \left\langle \frac{\nabla_r P|_{plant}}{\left\| \nabla_r P|_{plant} \right\|_2}, \frac{\nabla_r P|_{plant}}{\left\| \nabla_r P|_{plant} \right\|_2} \right\rangle$$

Subject to: (2.11)

$$\left(\nabla_r P|_{plant} - \overline{\nabla_r P|_{plant}} \right) \mathbf{Q}_{plant}^{-1} \left(\nabla_r P|_{plant} - \overline{\nabla_r P|_{plant}} \right)^T = c^2$$

$$\left(\nabla_r P|_{model} - \overline{\nabla_r P|_{model}} \right) \mathbf{Q}_{model}^{-1} \left(\nabla_r P|_{model} - \overline{\nabla_r P|_{model}} \right)^T = c^2$$

Since, the extremum two points lie on the surface of the contours, only the equality of the constraints is considered. Solution of this optimization problem will give the maximum angle between the two cones.

2.8 Performance Evaluation Using Gradients

Once the normalized inner product of the gradients and the plant-model mismatch angle are obtained, they can be used to predict the efficiency of the process moves. A simplified flowchart is used to illustrate this procedure (see Figure 2.6).

- If the inner product of the two gradients is greater than zero, then the process is generally moving in a direction of improving profit.
- On the other hand, if the inner product of the two gradients is less than zero, then the optimizer is moving the process in a direction of decreasing profit.

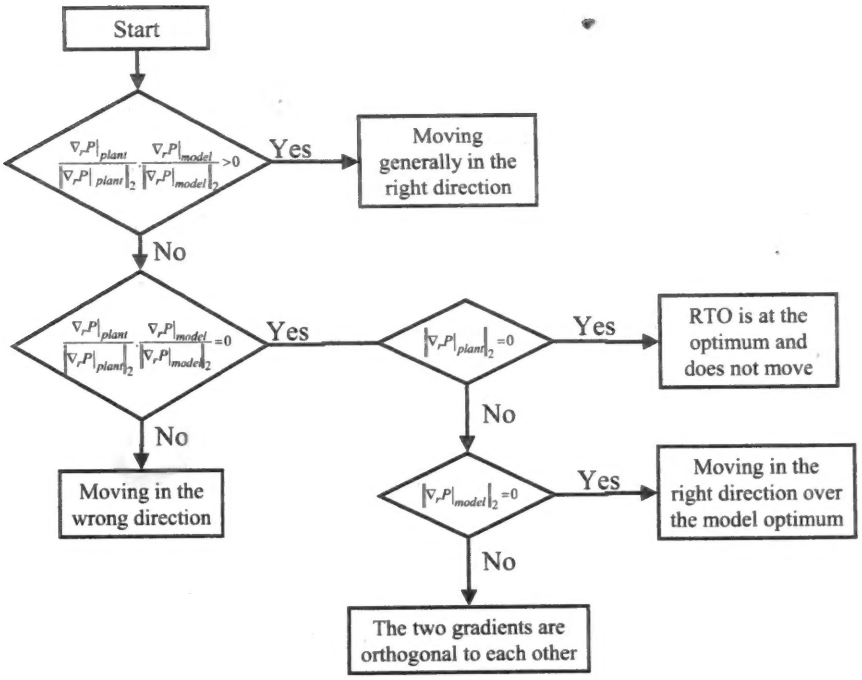
Start
$\frac{\nabla_r P|_{plant}}{|\nabla_r P|_{plant}|_2} \cdot \frac{\nabla_r P|_{model}}{|\nabla_r P|_{model}|_2} > 0$
Yes
Moving generally in the right direction
No
$\frac{\nabla_r P|_{plant}}{|\nabla_r P|_{plant}|_2} \cdot \frac{\nabla_r P|_{model}}{|\nabla_r P|_{model}|_2} = 0$
Yes
$|\nabla_r P|_{plant}|_2 = 0$
Yes
RTO is at the optimum and does not move
No
No
Moving in the wrong direction
$|\nabla_r P|_{model}|_2 = 0$
Yes
Moving in the right direction over the model optimum
No
The two gradients are orthogonal to each other

Figure 2.6: Flow Chart

- The inner product of the gradients can be zero under three conditions:

  1. *The plant gradient is zero*: This happens when the process is at the plant optimum. This is the point where the process should be operated. Since the RTO system only looks at the model gradient, the RTO system will gradually move the process away from this optimum unless both gradients are zero.

  2. *The model gradient is zero*: The optimizer considers the current operating point to be the process optimum and will drive the process to this point. Although the RTO system will maintain the operation at this point, but in presence of plant-model mismatch the plant gradient will not be zero.

  3. *The gradients are orthogonal to each other*: In this case, both gradients are

non-zero, but their inner product is zero. This means that the gradients are orthogonal or perpendicular to each other. In this case, the process will move along iso-profit lines if infinitesimally small moves are taken, but larger steps might drive the process to less profitable regions.

2.9 Plant-Model Mismatch Cases

It has been shown that the *plant-model mismatch angle* changes as the RTO setpoints approach optimum operating point. Monitoring the gradual change of the mismatch angle can also provide useful information about the location of the true optimum relative to the process operating point and the model optimum. Based on the relative orientation of the plant optimum, model optimum and the current operating point on the profit surface, three different scenarios can be observed. These scenarios are elaborated below:

1. ***Case I***: Consider a situation where the model optimum is located between the plant optimum and the current operating point on the profit surface. In such a case, relative to the current operating point, the plant and the model optimum are generally in the same direction (see Figure 2.7). Therefore, the profit gradients are also pointing at the same direction, and the normalized inner product of the gradient (Φ) is close to one. For such a condition, the performance of the RTO system is well within the expectations, and this is reflected by the mismatch angle which will be close to zero.

   The performance of the RTO system will start to change as the process operating conditions approach the model optimum. At the optimum for the model, the reduced profit gradient of the model ($\nabla_r P_{model}$) will be zero. Therefore, the inner product of the gradients, Φ will also be zero, even though the profit gradient of the plant will be much larger.

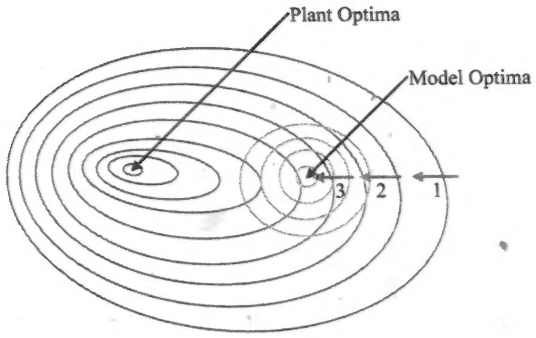

Figure 2.7: Model optimum between plant optimum and initial operating point

In summary, when the model optimum lies between the plant optimum and the current operating point, the mismatch angle may remain near zero until the process converges to the model optimum. After reaching the model optimum, the RTO system will most likely not progress further towards the plant optimum.

2. ***Case II***: In this case, consider the plant optimum to be between the model optimum and current operating point (see Figure 2.8).

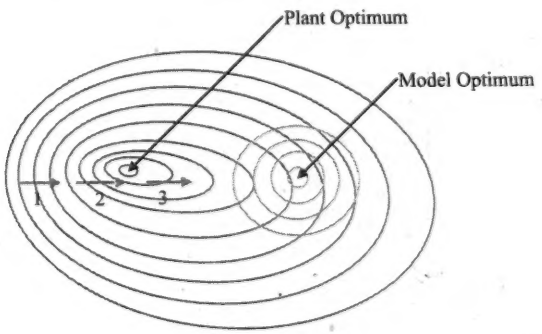

Figure 2.8: Plant optimum between model optimum and initial operating point

Similar to case Case *I*, at the initial operating point the profit gradients will be pointing generally in the same direction, and mismatch angle between them will be close to zero. As the RTO system drives the process towards plant optimum, the profit gradient of the plant ($\nabla_r P_{plant}$) will approach zero, and the normalized inner product of the gradients (Φ) at the plant optimum will also be zero. In this case, however, the model based RTO system will continue past the plant optimum towards the model optimum. Once the process crosses the plant optimum, the two optima will be on two opposite sides of the operating point. This will cause a sudden switch in sign of the inner product of the gradients from positive to negative.

The plant-model mismatch angle will be close to 180° for any operating point between the gradients. This will be a clear indication that the RTO is moving away from the actual optimum. When the RTO setpoints converge to the model optimum, the model gradients will be reduced to zero and the product of the inner gradients will again be zero.

3. ***Case III***: Case *III* can be considered as a more generalized form of Case *II*. In this case the initial operating point is at an angle with the line connecting the two optima and close to the plant optimum (See Figure 2.9).

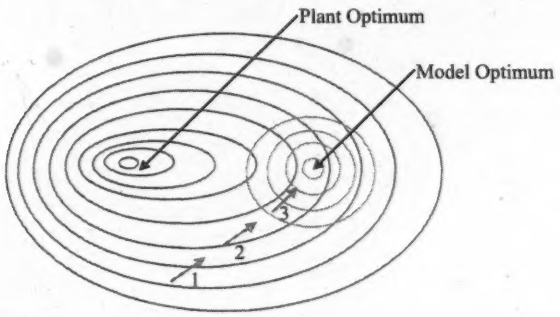

Figure 2.9: Model optimum between plant optimum and initial operating point

The operating point and the two optima are not aligned with each other, as in Case I or II. Thus the change in the gradient angle may not be as abrupt as Case *II*. The angle will be initially $< 90°$, and gradually it will change to $> 90°$ as is moves towards the model gradient and leaving the the plant gradient behind.

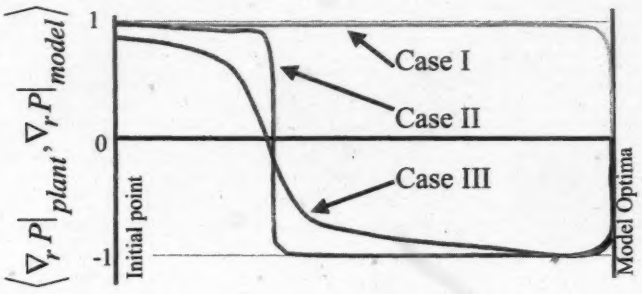

Figure 2.10: Changes in the product of the gradients

Based on this discussion it can be noted that changes in the inner product of the gradients can provide useful information on RTO performance. Figure 2.10 shows the change in the inner product of the gradients as the plant moves towards the model optimum (for all three cases discussed above). In all three cases, the inner product of the gradients are initially close two one. As the process moves towards its model optimum, the value of the inner product changes, and this change depends of the relative location of the plant optimum, the model optimum and the operating point. Therefore, in some cases, monitoring the value of the inner product can help estimate the direction of the true optimum.

2.10 Example Case Study

A simple three heat-exchanger network system has been used to illustrate the proposed plant-model mismatch detection techniques. A detailed description and schematic

diagram (see Figure A.2) of the model is given in Appendix **A**. The objective of this example is to demonstrate (a) the technique of RTO system performance evaluation using the plant-model mismatch angle, and (b) the gradual change of the mismatch angle as it approaches the model optimum. Two cases studies have been performed based on the initial operating point. They are described below:

CASE I

For this heat-exchanger network, the plant and model optima are at $[0.492, 0.167]$ and $[0.504, 0.168]$ respectively. In the first RTO step, the operating point is considered to be at $[0.300, 0.300]$. The location of the operating point with respect to the two optima can be compared to the situation explained in Case III of section 2.9. The setpoints implemented by the RTO system and their corresponding plant-model mismatch angles are given in Table 2.2.

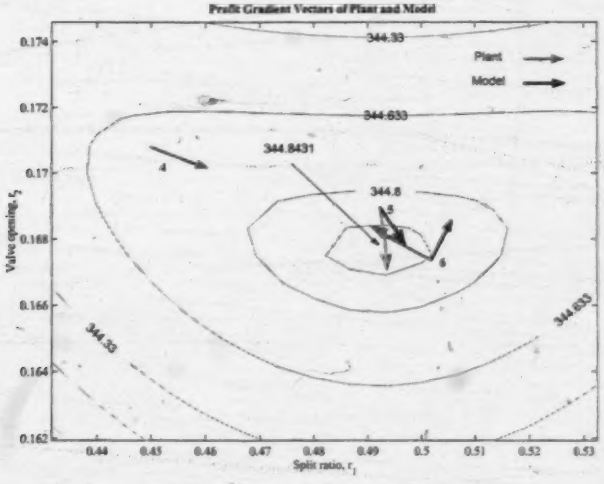

Figure 2.11: Gradient vectors on profit contour from 4^{th}, 5^{th} & 6^{th} RTO steps

In the first three RTO steps, the optimization routine takes the maximum allowed magnitude of the RTO steps, which is set at 0.05 for both manipulated variables.

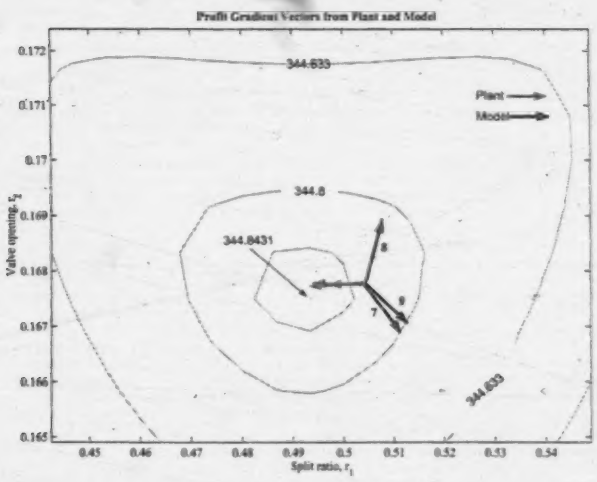

Figure 2.12: Gradient vectors on profit contour from 7^{th}, 8^{th} & 9^{th} RTO steps

The first five RTO steps quickly drive the process close to the optimum operating condition. The small mismatch angles for these five RTO steps indicate that the model is in the right direction. In the sixth RTO step, where the optimizer has crossed the model optimum, the angle between the gradients become quite large (i.e., 77.74°). This is an indication that the RTO system may be taking a wrong move. The fourth, fifth and sixth RTO steps are shown in Figure 2.11. The gradient vectors in the plot show the appropriate direction of the profit gradients, but their magnitudes have been modified, i.e., scaled to unity to fit the scale of the contour plot.

In the seventh RTO step, the optimizer reaches the model optimum. At this point, the model gradient is close to zero. Considering the model optimum to be the true optimum, the RTO system does not move the setpoints any further. In the eighth and the ninth RTO steps the direction of the profit gradients change but the magnitude of the gradient remain close to zero. Therefore, the RTO system reaches the stationary point at the model optimum.

CASE II

In this case the initial operating point ([0.533, 0.168]) is selected to be on the line connecting the points [0.492, 0.167] and [0.504, 0.168], which correspond to the plant and model optima respectively. Note that in this case the model optimum is between the plant optimum and the current operating point. The setpoints and mismatch angle for the RTO steps are given in Table 2.3.

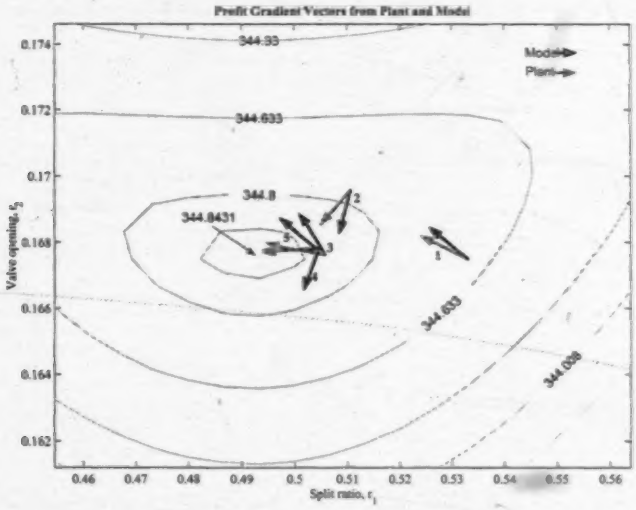

Figure 2.13: Gradient vectors on profit contour from the 1st to the 5th RTO steps

Figure 2.13 shows the reduced profit gradients of the plant and its model. It can be noted that in the first RTO step, the setpoints do not change in the direction of the model optimum. This is because the optimizer has taken action on the basis of the local gradient direction, which is orthogonal to the profit contours. In the first two RTO steps, the angle between the gradients are quite small because both these two points are away from the optimum. In the third RTO step, the process converges very close to the model optimum. In the rest of the RTO steps, the setpoints remain almost unchanged. It is evident from Figure 2.14 that the process is not moving any

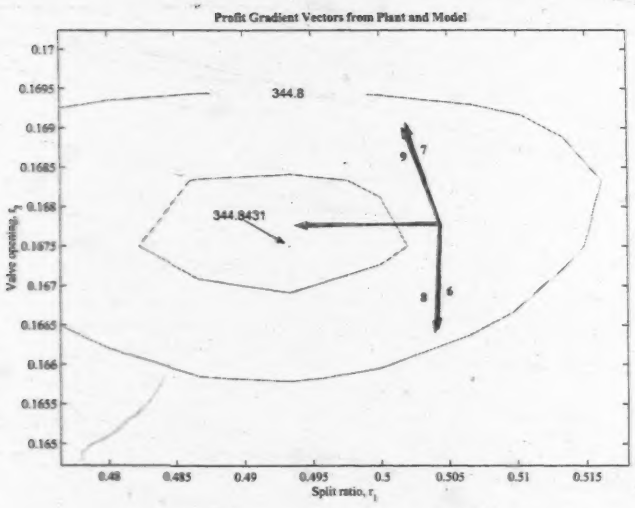

Figure 2.14: Gradient vectors on profit contour from the 6^{th} to the 9^{th} RTO steps

more from the model optimum and the mismatch angle for these steps remain close to 90°.

In summary it can be said that the comparing the profit gradients from the model and the plant can help diagnose the performance of the RTO system at any RTO setpoint of the process.

Table 2.2: RTO setpoints and their corresponding mismatch angles (Case I)

RTO Step No.	**Operating setpoints (split ratio)**	**Mismatch angle (degrees)**
1	[0.300, 0.300]	3.181
2	[0.350, 0.250]	2.365
3	[0.400, 0.200]	1.905
4	[0.450, 0.171]	0.178
5	[0.492, 0.169]	21.293
6	[0.502, 0.167]	77.740
7	[0.504, 0.168]	133.622
8	[0.504, 0.168]	109.358
9	[0.504, 0.168]	145.359

Table 2.3: Operating point and their corresponding mismatch angle (Case II)

RTO Step No.	Operating setpoints (split ratio)	Mismatch angle (degrees)
1	[0.533, 0.168]	13.646
2	[0.511, 0.170]	22.422
3	[0.506, 0.168]	52.857
4	[0.505, 0.168]	68.899
5	[0.504, 0.168]	46.963
6	[0.504, 0.168]	87.108
7	[0.504, 0.168]	74.162
8	[0.504, 0.168]	87.175
9	[0.504, 0.168]	75.907

Chapter 3

Diagnosis of Plant-Model Mismatch

The process industry started to implement online optimization projects in the early 1960s (Darby and White, 1988). These projects used correlation-based "engineering" process models or short-cut models to minimize computation time, which was a serious limitation of the computers at that time. Over the years, the advent of faster and cheaper computers has encouraged engineers to use more rigorous process models to take full advantage of the potential benefits of optimization (Sourander *et al.*, 1994). The number of applications has significantly increased in the last twenty years, but the success rate has not improved in the same proportion. Some of the reasons for these failures have been discussed in Chapter 1. Among these, one of the primary reasons has been the inability of the RTO system to ensure that it is truly taking the process towards increasing profitability. The complexity of the models and the associated calculations have made it impossible to look at the the RTO results (i.e., setpoints for the process), and draw conclusions about their correctness. Even if the model and its parameter updating schemes are working properly, the RTO system can still move the process away from its optimum operating conditions. This is because the RTO system uses process information from previous operating conditions to predict the new setpoints, where as the current process parameters might be different from the estimated parameters. Some of these mismatches will have no influence on the process performance, whereas some others will cause significant loss of productivity. It is necessary to identify the mismatches that are related to perfor-

mance loss and correct them. Currently, the only available method for performance evaluation is to go through the complete coding of the model and its updater, and find the discrepancy in the code causing the mismatch. This is a very expensive and time-consuming process, which requires skilled personnel from multiple disciplines on a long term basis. To avoid high costs, industry feels the need for easier RTO performance evaluation techniques. The techniques should be capable of identifying the root causes of performance loss, and at the same time be straight forward enough for the process operators to understand. In addition, combinations of mismatch can lead to profit loss, and techniques are required to identify these combinations.

Diagnosis is defined as "the art or act of investigation or analysis of the cause or nature of a condition, situation, or problem" (Merriam-Webster, 1994). For an RTO system, diagnosis would mean finding the cause of incorrect set-point predictions by the RTO system. In most cases, these incorrect estimations can be traced back to the process model or the parameter updating approach. It is not possible to obtain a model that will be an exact representation of the process and no parameter estimation technique can give perfect estimates of the parameters with finite amounts of data. That is why there will always be plant-model mismatch (see Chapter 2). This is especially true for nonlinear processes. Simplifying assumptions are made in the model to compensate for the nonlinearity, and this makes it vulnerable to setpoint changes away from its design condition. Identifying and rectifying these mismatches can be quite difficult and time consuming.

As mentioned before, some forms of plant-model mismatch will be more detrimental to the RTO system performance than others. It is necessary to prioritize the mismatch and investigate only the ones that are causing significant loss of performance. There should be a measuring scale, based on a selection or weighting criterion, to sort these plant-model mismatches according to their impact on process performance. For the majority of the optimization algorithms, the objective is to maximize productivity of the process, e.g., maximize profitability. Hence, it is logical to prioritize the mismatch based on contribution to productivity or profit loss. For a

control engineer, the task would be to find the opportunities to increase productivity by eliminating related plant-model mismatch from the RTO system.

In the current structure of RTO systems, the objective of the *optimization block* (see Chapter 1 for details) is to ensure that the RTO system is generating more profit in every step than its previous step. These set-point changes are estimated using the model and the profit function. The optimization algorithm calculates the change in profitability as the process moves over the profit surface, and predicts a direction of profit increase. In other words, it conducts a sensitivity study on the profit surface generated from the process model. The profit surface of the model has to match the plant profit surface, especially around the current operating point, to ensure accurate RTO prediction. To guarantee reasonable accuracy, the model parameters are re-calculated in every RTO cycle using available steady-state process data (see *parameter updating block* in Chapter 1). The *data reconciliation block* scrutinizes the regular operating data to select the disturbance-free, steady-state data for parameter estimation. It is evident from this discussion that the performance of each block in the RTO loop depends on the performance of its preceding blocks. For example, if the data reconciliation block fails to distinguish between process disturbance and normal operation, it will pass on the upset condition data for parameter estimation block. The parameters, estimated using data from process upset conditions, may not represent the process parameters accurately. When these inaccurate parameters are used in the model, it will give erroneous prediction of the process outputs. It is evident that any fault in the optimization block, the parameter updating block, or the data reconciliation block will contribute to incorrect estimation of the plant profit and drive the process away from its true optima. This clearly illustrates the complexity of isolating an error in the RTO system and justifies developing methods for detection of RTO performance degradation.

There have been attempts to tackle the plant-model mismatch problem from different perspectives of process control. There are papers that propose careful selection of the model (i.e., first-order model, black box model or hybrid model), and choosing

the one with minimum structural and parametric mismatch (Forbes, 1994). Others have attempted to improve model performance by improving parameter estimation techniques (Box, 1970; Krishnan *et al.*, 1992). These methods of better model selection and parameter updating are valid for cases where the RTO system is under development and has not yet been implemented. Once the RTO system has been put online, it is not possible to change the whole model or its parameter updating technique without shutting down the system. A convenient alternative is to upgrade the section of the model that is contributing to performance loss. The first step for this is to identify that problematic part of the model or the parameter updating block.

In this thesis, tools are proposed which allow RTO users to isolate the section of the model contributing to significant performance or profit loss.

3.1 Tool Selection

The objective for the remainder of this chapter is to propose new tools for plant-model mismatch analysis. A logical first step in mismatch analysis is to compare the predictions from the model with the actual plant measurements, and quantify the prediction error. The process measurements can be represented in different ways, and based on that, the technique of comparing the plant-model mismatches can also be quite different. One method might compare the individual sensor reading (i.e., the measured variables) with their corresponding estimates from the model. This will be helpful in understanding the characteristics of mean, standard distribution, skewness of that variable, but will not help in understanding the interaction between the manipulated and measured variables of the process. Another method might look at the variance-covariance matrix of the process outputs and compare them with the variance-covariance of the model outputs. Comparing the variance-covariance matrix is a popular method of mismatch analysis in multivariate statistics, like fault detection. Fault detection compares the variance-covariance matrix of the process

data from normal conditions and upset conditions to identify the source of the fault. This type of analysis technique (e.g., *Principle Component Analysis* (PCA)) can be applied only to mean-centered and scaled data matrix (Golub and Loan, 1990). Scaling eliminates the opportunity to study the magnitude of the mismatch, which is an important part of plant-model mismatch diagnosis. *Singular Value Decomposition* (SVD) is the general form of PCA. SVD does not anticipate any kind of scaling of the data matrix and it is sensitive to to scaling of the data. This thesis uses a modified form of SVD to identify the source of the plant-model mismatch.

As explained in Chapter 2, the sensitivity matrix ($\frac{d\mathbf{y}}{d\mathbf{x}}$) of a process shows the sensitivity of the measured variables ($\mathbf{y}$) to changes in the manipulated variables ($\mathbf{x}$). The sensitivity matrix of the input-output variables presents the interaction of the variables in a compact matrix format, which makes it convenient for mathematical analysis. Each of the columns of the sensitivity matrix corresponds to one of the manipulated variables, and each of the rows are related to one of the measured outputs:

$$\underbrace{\begin{bmatrix} \partial y_1 \\ \partial y_2 \\ \vdots \\ \partial y_m \end{bmatrix}}_{Output} = \underbrace{\begin{bmatrix} \frac{dy_1}{dx_1} & \frac{dy_1}{dx_2} & \cdots & \frac{dy_1}{dx_n} \\ \frac{dy_2}{dx_1} & \frac{dy_2}{dx_2} & \cdots & \frac{dy_2}{dx_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{dy_m}{dx_1} & \frac{dy_m}{dx_2} & \cdots & \frac{dy_m}{dx_n} \end{bmatrix}}_{Sensitivity\ matrix} \underbrace{\begin{bmatrix} \partial x_1 \\ \partial x_2 \\ \vdots \\ \partial x_n \end{bmatrix}}_{Input} \tag{3.1}$$

The elements of the first column of the sensitivity matrix, i.e., $[\frac{dy_1}{dx_1}, \frac{dy_2}{dx_1}, \ldots, \frac{dy_m}{dx_1}]'$, are related to the first manipulated variable, x_1. Similarly, the other columns of the matrix are associated with x_2, x_3, and so forth. On the other hand, the first row of the matrix, i.e., $[\frac{dy_1}{dx_1}, \frac{dy_1}{dx_2}, \cdots, \frac{dy_1}{dx_n}]$, describe the sensitivity effects for the first measured variable y_1, and so on. For a perfect model of a process, each of the elements of the model sensitivity matrix will match the corresponding element of the plant sensitivity matrix. If one of the elements (for example $\frac{dy_i}{dx_j}$) in these two sensitivity matrices does not match, this indicates that the model has failed to predict the change in the output y_i of the process in response to a change in the input x_j. This hints at one of two faults in the RTO system:

1. in the model, the equations relating the input (x_j) and the output (y_i) are inappropriate for that operating condition, or

2. the parameters related to that section of the model have not been correctly updated in the model-updater block.

For both these cases, comparing the sensitivity matrices can help identify the equations related to a mismatch. For the rest of the thesis, the difference between the plant and model sensitivity matrices will be referred to as the *error matrix*, **E**:

$$\mathbf{E} = \left(\frac{d\mathbf{y}}{d\mathbf{x}}\bigg|_{plant} - \frac{d\mathbf{y}}{d\mathbf{x}}\bigg|_{model}\right) \tag{3.2}$$

For the sections of the process that are exactly represented by the model, the elements of the error matrix **E** may be zero. On the other hand, where the process is misrepresented by the model, the elements of **E** may be non-zero. If only one section of the model is incorrect (due to structural or parametric plant-model mismatch), it may be distinguished from the rest of the model by scrutinizing the non-zero elements of **E** related to that section. However, due to excessive process noise, estimates of $\frac{d\mathbf{y}}{d\mathbf{x}}|_{plant}$ and $\frac{d\mathbf{y}}{d\mathbf{x}}|_{model}$ can sometimes be less accurate than desired. This adds extra non-zero and zero terms in the error matrix, **E**, making it difficult to segregate the true mismatch from the effect of noise. This thesis assumes that the process noise has zero mean and Gaussian distribution.

The units and scale of the elements in the **E** matrix are usually different. The magnitude of one element can be much larger than another element. Entries in the error matrix (**E**) will have units of $[Output\ Unit]/[Input\ Unit]$. A small mismatch in a large gradient can show up more prominently in **E**, than a large mismatch in a smaller gradient. This makes it difficult to study the error matrix (**E**) and identify the true mismatch. For example, a small mismatch in product concentration might be more important than a larger mismatch in cooling water flowrate. Therefore, it is necessary to present all the mismatches in a consistent scale or unit.

As discussed in chapter 2, the reduced profit gradient ($\nabla_r P$) represents the sensitivity of the profit function to the manipulated variables ($\mathbf{x}$). Changes in the manipulated variables ($\mathbf{x}$) modify the profit gradient ($\nabla_r P$) both explicitly (via $\frac{\partial P}{\partial \mathbf{x}}$) and implicitly (via the process outputs $\mathbf{y}$), which changes the plant sensitivity matrix, $\frac{d\mathbf{y}}{d\mathbf{x}}$ (see Equation 2.4). This equation (Equation 2.4) takes the sensitivity matrix ($\frac{d\mathbf{y}}{d\mathbf{x}}$), and presents it as the sensitivity of the plant profit function ($P(\mathbf{x}, \mathbf{y}, \beta)$) with respect to the manipulated variables ($\mathbf{x}$). The gradient of the profit function with respect to the outputs ($\frac{\partial P}{\partial \mathbf{y}}$) can be considered as a weighting vector that transforms the sensitivity matrix into its contribution in the reduced profit gradient ($\nabla_r P$).

The error matrix ($\mathbf{E}$) has the same dimensions and units as the sensitivity matrix ($\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$) whose impact on profit is weighted by $\frac{\partial P}{\partial \mathbf{y}}$. It is beneficial to examine the error matrix scaled by the profit gradient with respect to the outputs $\frac{\partial P}{\partial \mathbf{y}}$ as well. The product of the profit gradient with respect to the outputs ($\frac{\partial P}{\partial \mathbf{y}}$) and the error matrix ($\mathbf{E}$) gives the amount of lost profit with respect to each of the manipulated variables. Defining the product as *lost profit* $\mathbf{\Delta P}$ gives,

$$\Delta P \equiv \frac{\partial P}{\partial y}\mathbf{E} \tag{3.3}$$

The quantity ΔP in Equation 3.3 is a vector identifying the impact of the mismatch on profit according to each input. Instead of multiplying the vector and the matrix, efforts should be made to study their interaction and their influence on each other, thereby identifying mismatches based on their influence on profitability.

Based on the assumption that the price of the raw materials and utilities are exactly known (see Assumptions in Chapter 1), the impact of the manipulated variables ($\mathbf{x}$) and the measured variables ($\mathbf{y}$) on the profit function (P) can be exactly calculated for both the plant and the model. In other words, $\frac{\partial P}{\partial \mathbf{y}}\Big|_{plant} = \frac{\partial P}{\partial \mathbf{y}}\Big|_{model}$ and $\frac{\partial P}{\partial \mathbf{x}}|_{plant} = \frac{\partial P}{\partial \mathbf{x}}|_{model}$. If these conditions are satisfied, ΔP in Equation 3.3 corresponds to the subtraction of the reduced profit gradients (Equation 2.4 and 2.5):

$$\begin{aligned}\Delta P &\equiv \left(\nabla_r P|_{plant} - \nabla_r P|_{model}\right) \\ &\doteq \frac{\partial P}{\partial \mathbf{y}}\left(\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} - \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model}\right) \\ &= \frac{\partial P}{\partial \mathbf{y}}\mathbf{E}\end{aligned} \tag{3.4}$$

In equation 3.4, $\Delta P = [p_1, p_2, ..., p_r]$, where r is the number of manipulated variables of the process. The p_i's are the sensitivity of profit to the model mismatch. The manipulated variable corresponding to the largest p_i is related to the section of the model that is contributing to the maximum amount of profit loss. Searching through all the equations that are related to this manipulated variable might help find the cause of this mismatch, but it will be quite time consuming. It is necessary to get better idea about the root cause of the mismatch before trying to solve the problem.

Equation 3.4 represents the lost profit as the product of a vector ($\frac{\partial P}{\partial \mathbf{y}}$) and a matrix ($E$). In particular, the ith element of ΔP represents the inner product of $\frac{\partial P}{\partial y}$ and the ith mismatch (column) vector, $\mathbf{e}_i$, in $\mathbf{E}$. Investigating the interaction of the two can provide useful information about the plant-model mismatch and can be accomplished using matrix decomposition techniques. Among different methods of matrix decomposition, two approaches have been found feasible for mismatch study:

1. QR decomposition is used to redefine the error matrix with a new set of basis vectors, one of which is the profit gradient. This technique is useful to study the orientation of columns of the error matrix with respect to the profit gradient, i.e., the relative orientation of mismatch sensitivity for a given input relative to the profit gradient, $\frac{\partial P}{\partial \mathbf{y}}$.

2. SVD decomposition is done to a modified error matrix, where all the elements of the matrix have been weighted according to their profit function sensitivity. This technique can break the error matrix down into similar-sized matrices of one range space and one domain space.

Other methods of matrix decomposition and analysis were also explored, which were not found to be suitable for diagnostic purposes. Some of them are briefly discussed below:

- Principle Component Analysis (PCA) is a special case of SVD decomposition. In plant-model mismatch detection it does not provide any additional information when compared to the information gathered from SVD decomposition.
- LU decomposition can only be applied to square matrices. The sensitivity matrix of the process is usually non-square, because processes have more measured variables than manipulated variables. This is why LU decomposition is not a viable option for mismatch diagnosis.

The proposed two methods for mismatch diagnosis - (a) modified QR decomposition and (b) SVD decomposition, are elaborated in the following sections.

3.2 Diagnosis using Modified QR Decomposition

Orthogonal matrix triangularization, also known as *QR decomposition*, takes the linear mapping of a vector space from one basis and maps it using a new orthonormal basis for the range space (Golub and Loan, 1990). The QR decomposition of matrix $\mathbf{A}$ (where, $\mathbf{A} \in \Re^{m \times n}$) has the form:

$$\mathbf{A} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \tag{3.5}$$

The columns of the orthogonal matrix $\mathbf{Q}$ (i.e., $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$) are the new basis vectors of the range space. The first p columns (where, $p = rank(\mathbf{A})$) of $\mathbf{Q}$ span the range space of $\mathbf{A}$, and the remaining columns of $\mathbf{Q}$ are orthogonal to the range space of $\mathbf{A}$.

In the following subsection, QR decomposition will be used as a tool to investigate the orientation of the error matrix ($\mathbf{E}$) with respect to the profit gradient.

3.2.1 Diagnostic Tools from Modified QR

It is evident from Equation 3.4 that the orientation of the columns of the error matrix, **E** with respect to the profit gradient determines the magnitude of the lost profit. If the column vectors of **E** are in the same direction as the profit gradient, the profit loss will be high; whereas if the columns of **E** are nearly orthogonal the profit loss will be minimal.

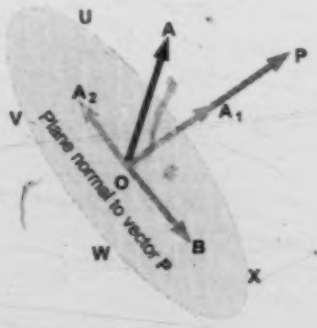

Figure 3.1: Influence of vector alignment with profit gradient $\frac{\partial P}{\partial y}$ on the overall profit

The effect of the orientation of the vectors with respect to profit loss is illustrated in Figure 3.1. OP is the profit gradient and OA and OB are two vectors of the error matrix. A new set of basis vectors can be used to represent these vectors. If the vectors are in an n-dimensional space, n new basis vectors will be needed to span the complete space. Defining the profit gradient, OP as one of the new basis vectors, then using QR decomposition, the remaining $n-1$ basis vectors will be orthogonal to OP and lie on the $UVWX$ plane (where vector OP is perpendicular to the $UVWX$ plane). In the new co-ordinate system, vector OA is split into two components OA_1 and OA_2. OA_1 is aligned with OP, and OA_2 is on the orthogonal plane $UVWX$. OA_2 can be represented as a linear combination of the $n-1$ basis vectors spanning the $UVWX$ plane. Similarly, the vector OB can be represented by two new vectors,

where one component is parallel to OP and the other component is orthogonal to OP.

The dot product of vector OP and OA will only be influenced by the component OA_1. The orthogonal component, OA_2 will have no influence on the product, as shown below:

$$\begin{aligned}
&\langle OP, OA \rangle \\
&= \langle OP, OA_1 \rangle + \langle OP, OA_2 \rangle \\
&= \|OP\|_2 \cdot \|OA_1\|_2 \cos \angle A_1OP + \|OP\|_2 \cdot \|OA_2\|_2 \cos \angle A_2OP \\
&= \|OP\|_2 \cdot \|OA_1\|_2 \cos 0° + \|OP\|_2 \cdot \|OA_2\|_2 \cos 90° \qquad (3.6) \\
&[\because \text{OP and } OA_2 \text{ are orthogonal to each other by definition}] \\
&= \|OP\|_2 \cdot \|OA_1\|_2 \\
&= \langle OP, OA_1 \rangle
\end{aligned}$$

In this example, OA_2 is orthogonal to OP, and hence its dot product will be zero in the new basis.

It is evident that only the vector components parallel to the profit gradient ($\frac{\partial P}{\partial y}$) contribute to profit loss. Any vector or any component of a vector perpendicular to the profit gradient will have no impact of the profit function. This is an important property of the vectors for mismatch identification and diagnosis. Redefining the vector space of the $\mathbf{E}$ matrix using a new set of basis vectors, to show the orientation of the columns of $\mathbf{E}$ with respect to the profit gradient can provide useful information about the mismatch. QR decomposition is used as a tool to transform the representation of the original error matrix $\mathbf{E}$ to a new set of basis vectors, where one of the basis vectors is parallel to the profit gradient ($\frac{\partial P}{\partial y}$).

As mentioned earlier, QR decomposition is a method for reorganizing the basis vectors of a matrix system. In this decomposition technique, the first column $\mathbf{a}_1$ of a given matrix $\mathbf{A}$ (where $\mathbf{A} \in \Re^{m \times n}$ and $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$) is taken as the

first new basis vector, $\mathbf{q}_1$ (where $\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$) after normalizing it, i.e, $\mathbf{q}_1 = norm(\mathbf{a}_1)$ (Golub and Loan, 1990). The second basis vector $\mathbf{q}_2$ is created by taking the component of the $\mathbf{a}_2$ vector that is orthogonal to the first basis vector, and normalizing. Similarly, the other basis vectors are generated from the component of each vector that is orthogonal to all the basis vectors before it, followed by normalization. This can be accomplished by Gram-Schmidt orthogonalization or the use of Householder transformation (Golub and Loan, 1990).

The matrix $\mathbf{R}$ is the representation of the mapping of $\mathbf{A}$ in the new orthogonal basis, $\mathbf{Q}$. The first p columns of $\mathbf{Q}$ (where $p = \min(m, n)$) span the range space of the matrix $\mathbf{A}$, while the remaining columns are orthogonal to the range space.

QR decomposition of a matrix in which the profit gradient ($\frac{\partial P}{\partial \mathbf{y}}$) is the first column will create a matrix of new basis vectors ($\mathbf{Q}$), where the first basis vector ($\mathbf{q}_1$) will be in the direction of the profit gradient, and the remainder orthogonal to that direction. This is done by augmenting the error matrix ($\mathbf{E}$) by inserting the profit gradient ($\frac{\partial P}{\partial \mathbf{y}}$) in the first column, and performing QR decomposition of this augmented matrix.

Given the profit gradient is $\frac{\partial P}{\partial \mathbf{y}} = \begin{bmatrix} \frac{\partial P}{\partial y_1} & \frac{\partial P}{\partial y_2} & \cdots & \frac{\partial P}{\partial y_m} \end{bmatrix}$ and the error matrix is:

$$\mathbf{E} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{bmatrix} \tag{3.7}$$

The new augmented matrix for QR decomposition is:

$$\tilde{\mathbf{E}} = \begin{bmatrix} \frac{\partial P}{\partial y_1} & e_{11} & e_{12} & \cdots & e_{1n} \\ \frac{\partial P}{\partial y_2} & e_{21} & e_{22} & \cdots & e_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial P}{\partial y_m} & e_{m1} & e_{m2} & \cdots & e_{mn} \end{bmatrix} \tag{3.8}$$

QR decomposition of the $\tilde{\mathbf{E}}$ matrix generates the basis matrix, $\mathbf{Q}$:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \end{bmatrix} \tag{3.9}$$

and the upper triangular matrix, $\mathbf{R}$:

$$\mathbf{R} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & \cdots & r_{1,(n+1)} \\ 0 & r_{2,2} & r_{2,3} & \cdots & r_{2,(n+1)} \\ 0 & 0 & r_{3,3} & \cdots & r_{3,(n+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & r_{m,(n+1)} \end{bmatrix} \tag{3.10}$$

where m and $(n+1)$ are the number of rows and columns respectively in the $\tilde{\mathbf{E}}$ matrix. The first column of $\mathbf{Q}$, $\mathbf{q}_1$ will be the normalized form of the profit gradient. The remaining columns, $\mathbf{q}_2, \mathbf{q}_3, \ldots, \mathbf{q}_n$ are the new set of basis vectors, which are orthogonal to the profit gradient ($\mathbf{q}_1$). The first row of the $\mathbf{R}$ matrix, $\begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1(n+1)} \end{bmatrix}$ represents the components of the column vectors of $\tilde{\mathbf{E}}$ (i.e., r_{12}, r_{13}, ..., $\mathbf{r}_{1(n+1)}$) that lie along the new basis vector q_1. The largest r_{1i}, where $i = 2, 3, .., m$ will be the biggest contributor to the profit loss. The $\mathbf{q}_i$ vectors corresponding to the largest r_{1i} (i.e., q_2 corresponds to r_{12}, q_3 corresponds to r_{13} and so forth) will be the manipulated variable with the largest impact on the profit loss.

3.2.2 Illustrative Example

A simple industrial process example is used to demonstrate the application of proposed modified QR decomposition technique. The process flowsheet is given in Figure 3.2. The process consists of a reactor and a distillation column. The pure reactant stream (F) is fed into the jacketed reactor, where a fraction of the reactant is converted to the product. Exothermic isomerization of reactant A produces the product B and heat. Since the reaction is exothermic, removal of heat from the reactor increases product concentration (x_i) in the stream leaving the reactor. Cold water (W) is used inside the reactor cooling jacket to regulate the reactor temperature. The mixed stream of product (B) and reactant (A) leaving the reactor is fed into the distillation column, which has a total condenser and a partial reboiler. The reflux rate of the distillation column is manipulated by changing the flowrate of product stream

(P). The objective for the optimization routine is to maximize the product flowrate (P) and purity (x_D), by regulating feed rate (F), cooling water flow-rate (W) and reboiler steam flow-rate (S).

The overall process has three manipulated variables, (1) Feed rate, F, (2) Cooling water rate, W, and (3) Steam flow rate to the reboiler, S. The measured variables for the process are: (1) mass fraction of product coming out of the reactor, x_i, (2) final product flow-rate, P and (3) mass fraction of product in the final product stream, x_D. Therefore, the sensitivity matrix of the process is a 3×3 matrix, where the measured variables, i.e., $\mathbf{x} = \begin{bmatrix} F & W & S \end{bmatrix}'$ are related to the manipulated variables, i.e., $\mathbf{y} = \begin{bmatrix} x_i & P & x_D \end{bmatrix}'$ by the following equations:

$$x_i = -0.096886 + 0.016888F - 0.000066F^2 + 0.025269W - 0.00035W^2 \tag{3.11}$$

$$\begin{aligned} P = -983.656 + 19.2599F - .089333F^2 + 3.8666W - 0.055555W^2 \\ -4.02357S + 0.063338S^2 \end{aligned} \tag{3.12}$$

$$\begin{aligned} x_D = .741788 + \ 0.03190F - 0.000160618F^2 + 0.170096W - 0.00246485W^2 \\ -0.326281S + 0.00616S^2 \end{aligned} \tag{3.13}$$

CASE I

In this case study two mismatches were introduced in the model. It was assumed that the following two parameters were inaccurately estimated:

- The reaction rate constant was 2% less than the actual value used in the plant.
- The heat of reaction was 5% more than the actual value.

The plant sensitivity matrix, estimated by perturbing the plant, and the model sensitivity matrix were found to be the following:

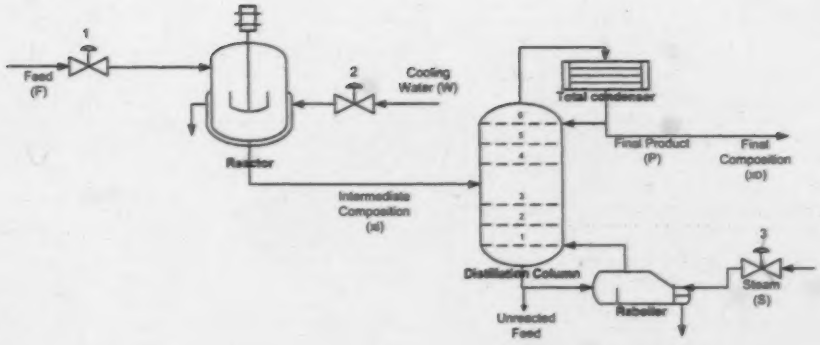

Figure 3.2: Illustrative Example of an Industrial Process

$$\left.\frac{dy}{dx}\right|_{plant} = \begin{bmatrix} 0.00325 & 0.00257 & 0.00000 \\ 0.51400 & 0.23200 & -0.66667 \\ -0.00094 & 0.000267 & 0.01480 \end{bmatrix}$$

$$\left.\frac{dy}{dx}\right|_{model} = \begin{bmatrix} 0.00345 & 0.00257 & 0.00000 \\ 0.52000 & 0.24230 & -0.66680 \\ -0.00744 & 0.00027 & 0.01440 \end{bmatrix}$$

Visually comparing the sensitivity matrices provide very little information about the plant-model mismatch. Most of the elements of the two matrices are relatively close and it is not possible to categorize the mismatches. It can be noted that the third element of the first row of both the matrices are zero, indicating no interaction between the corresponding input and output variables (intermediate composition, x_i does not depend on stream flowrate in the reboiler, S). The error matrix of the system is,

$$\mathbf{E} = \begin{bmatrix} 0.0002 & 0 & 0 \\ 0.0060 & 0.0103 & 0.0002 \\ -0.0065 & 0 & -0.0004 \end{bmatrix}$$

The error matrix indicates that the E_{22} has the largest mismatch, followed by

two mismatches of similar magnitude at E_{21} and E_{31}. The smaller plant-model mismatches in E_{11}, E_{22}, E_{23}, and E_{44} may be attributed to process noise, which have not been filtered by the data validation block. The profit gradient is determined to be $\frac{\partial P}{\partial \mathbf{y}} = \begin{bmatrix} 56.2 & 45.8 & 313.5 \end{bmatrix}$ for both the plant and the model.

The augmented matrix for the system is,

$$\left[\frac{\partial P}{\partial \mathbf{y}} \ \vdots \ \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\Big|_{plant} - \frac{\partial \mathbf{y}}{\partial \mathbf{x}}\Big|_{model} \right) \right] = \begin{bmatrix} 56.2 & 0.0002 & 0 & 0 \\ 45.8 & 0.0060 & 0.0103 & 0.0002 \\ 313.5 & -0.0065 & 0 & -0.0004 \end{bmatrix}$$

Modified QR decomposition of the augmented matrix gives:

$$Q = \begin{bmatrix} -0.1747 & -0.1650 & -0.9707 \\ -0.1423 & -0.9713 & 0.1907 \\ -0.9743 & 0.1715 & 0.1462 \end{bmatrix}$$

$$R = \begin{bmatrix} -321.7737 & 0.0054 & -0.0015 & 0.0004 \\ 0 & -0.0070 & -0.0100 & -0.0003 \\ 0 & 0 & 0.0020 & 0 \end{bmatrix}$$

The first row elements of $\mathbf{R}$, i.e., R_{1i}, where $i = 2, 3, ..., n$, quantify the orientation of the columns of $\mathbf{E}$ with respect to the profit gradient, $\frac{\partial P}{\partial y}$. The extent of profit loss caused by the columns of the error matrix depends on the orientation of the columns with respect to the profit gradient. In this case, $|R_{12}| > |R_{13}| > |R_{14}|$ (since $|0.0054| > |-0.0015| > |0.0004|$) indicating that the maximum profit loss is related to the first manipulated variable of the process (the feed flow rate, F) followed by the second manipulated variable (cooling water flowrate, W). This is contrary to what is observed from the error matrix, where E_{22} points to the largest mismatch. This shows that the largest plant-model mismatch in the error matrix (in this case E_{22}) may not always be the largest contributor to profit loss. Investigating the first column of $\mathbf{Q}$ shows that the largest element is the third element of the column, i.e., $Q_{31} = -0.9743$. The conclusion is that the largest mismatch is related with the first input (feed flowrate, F) and the third output (product composition, x_D) of the process.

In this case study, the modified QR decomposition technique identified the plant-model mismatches corresponding to the process feed flowrate (F) as the main cause of profit loss. This mismatch can be linked with the inaccurate reaction rate constant. On the other hand, the mismatch related to the cooling water flowrate (W), which is the second most significant cause of profit loss, can be attributed to the inaccurate heat of reaction. Even though the modified QR decomposition technique is detecting the mismatches and relating them with the manipulated variables, it is difficult to pinpoint the source of the mismatch without prior knowledge about the possible cause of the mismatch. Therefore, it will be necessary to have thorough understanding of the process and possible sources of plant-model mismatch for successful application of this technique.

CASE II

In this case study, the same process is used with a different cause of plant-model mismatch. It is assumed that the major mismatch is in the distillation column. The relative volatility of the product and the reactant, used in the distillation column model, has been incorrectly estimated. The sensitivity matrix of the model at its current operating conditions as follows:

$$\left.\frac{dy}{dx}\right|_{model} = \begin{bmatrix} 0.00322 & 0.00258 & 0.000 \\ 0.5138 & 0.2324 & -0.665 \\ -0.000122 & 0.000233 & 0.0152 \end{bmatrix}$$

The sensitivity matrix of the plant remains same as in Case **I**. The error matrix is given below:

$$\mathbf{E} = \begin{bmatrix} 0.00003 & -0.00001 & 0.00000 \\ 0.00020 & -0.00040 & -0.00200 \\ -0.000818 & 0.000034 & -0.00040 \end{bmatrix}$$

Modified QR decomposition of the error matrix produces the following **Q** and **R** matrices:

$$\mathbf{Q} = \begin{bmatrix} -0.1747 & -0.4574 & -0.8719 \\ -0.1423 & -0.8645 & 0.4820 \\ -0.9743 & 0.2083 & 0.0859 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} -321.7737 & 0.0008 & 0.0000 & 0.0007 \\ 0 & -0.0004 & 0.0004 & 0.0016 \\ 0 & 0 & -0.0002 & -0.0010 \end{bmatrix}$$

The R matrix indicates that there are two significant plant-model mismatches of similar magnitude related to the feed flowrate (F) and the steam flowrate (S) (see R_{12} and R_{14}). These two manipulated variables are related to two units of the process, i.e., the reactor and the distillation column. The inaccurate relative volatility in the distillation column is directly relating the profit loss with the steam flowrate (S). On the other hand, the feed flowrate (F), which is upstream from the distillation column, influences the final product flowrate (P) and product composition (x_D). Since P and x_D are incorrectly predicted by the model, so profit loss is also attributed to the feed flowrate (F). Conversely, the impact of cooling water on x_D is not significantly influenced by changes in relative volatility.

3.3 Diagnosis using Singular Value Decomposition

Singular value decomposition (SVD) is a general linear-algebraic tool for dealing with singular or near-singular matrices (Noble and Daniel, 1977). It is a generalization of eigenfunction analysis which decomposes an $m \times n$ matrix into three matrices mapping from one n dimensional space to an m dimensional space. It provides a wealth of information to the analyst, including orthogonal bases for the domain and range spaces, the 2-norm and Frobenius norm of the matrix, and a determination of the numerical rank of a matrix (Campbell, 1977). It is used extensively to decompose sparse matrices where other techniques like LU decomposition fail to give satisfactory results (Golub and Loan, 1990) . In this chapter, SVD has been used as a tool to

decompose the *error matrix* to isolate the section of the model causing maximum amount of profit loss.

The SVD of a real m-by-n matrix $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}' \tag{3.14}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices, and $\boldsymbol{\Sigma} = diag(\sigma_1, \sigma_2, ..., \sigma_r)$, $r = \min(m, n)$, with $\sigma_1 \geqslant \sigma_2 \geqslant ..... \geqslant \sigma_r \geqslant 0$. The σ_i are called the *singular values*. The singular values (σ_i) are the positive square roots of the eigenvalues of the $A'A$ matrix, i.e., $\sigma_i = \sqrt{\lambda_i(A'A)}$. The columns of $\mathbf{V}$ and $\mathbf{U}$ are referred as the *right singular vectors* and the *left singular vectors* respectively (Press *et al.*, 1977).

It is necessary to elaborate the concept of *nullspace* and *range* to illustrate the use of SVD for the RTO diagnosis purpose (Campbell, 1977; Golub and Loan, 1990). Consider a set of simultaneous equations

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3.15}$$

where $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{b}$ is an m dimensional vector in $\Re^m$, and $\mathbf{x}$ is an n dimensional vector in $\Re^n$. The equation defines $\mathbf{A}$ as a linear mapping from the vector space $\Re^n$ to the new vector space $\Re^m$. The *nullspace* is a subspace of $\Re^n$ and is the space of vectors mapped to zero, i.e., vectors $\mathbf{x}$ for which $\mathbf{A}\mathbf{x} = \mathbf{0}$. The dimension of the nullspace is a subspace of $\Re^n$ that is mapped to zero, and its dimension is called the *nullity* of $\mathbf{A}$.

There is also some subspace of $\Re(\mathbf{A})$ of $\Re^m$ that can be "reached" by $\mathbf{A}$, i.e., there exists some subspace of $\Re^n$ which is mapped to a subspace of $\Re^n$. This subspace of $\Re^m$ is called the *range* of $\mathbf{A}$. The dimension of the range is called the *rank* of $\mathbf{A}$. If $\mathbf{A}$ is square and nonsingular, its rank will be equal to the dimension of $\Re^m$ i.e., m. If $\mathbf{A}$ is square and singular, the rank will be less than m. In fact, the relevant theorem is, "*rank plus nullity equals m, where A is an $m \times m$ square matrix.*" (Golub and Loan, 1990).

SVD explicitly constructs orthonormal bases for the domain including the nullspace and range of a matrix. Specifically, the columns of **U** are an orthonormal set of basis vectors that span the range space, whereas the columns of **V** are an orthonormal basis for the domain of the mapping. The null space is contained in the basis vectors, represented by the columns of **V**, which correspond to the zero singular values ($\sigma_i = 0$) in $\mathbf{\Sigma}$.

The concept of the range space and null space for SVD decomposition can be explained with a simple example. Consider the following matrix **A** (See Equation 3.16).

$$M = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.16}$$

M is a 3×3 matrix, mapping a 3 dimensional space into a 3 dimensional space ($A : \Re^3 \rightarrow \Re^3$). It is a square matrix, and hence $rank + nullity = 3$. The third column of **M** is the sum of the first two columns, i.e., the matrix is singular and has rank 2. Consequently, the null space has dimension 1.

Singular value decomposition of **A** creates the three matrices U, $\mathbf{\Sigma}$ and V as given here after:

$$\begin{aligned} U &= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{\Sigma} &= \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ V &= \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \end{bmatrix} \end{aligned} \tag{3.17}$$

According to the previous discussion, the columns of $\mathbf{U}$ span the range space of $\mathbf{A}$. The diagonal values of $\mathbf{\Sigma}$ show that, the first two singular values are positive definite ($\sigma_1 = \sqrt{3}$, $\sigma_2 = 1$) but the third one is zero ($\sigma_3 = 0$). Therefore, $\mathbf{u}_1$ and $\mathbf{u}_2$ span the range of $\mathbf{A}$. This could have been anticipated, because in this case $nullity = 3 - rank = 3 - 2 = 1$.

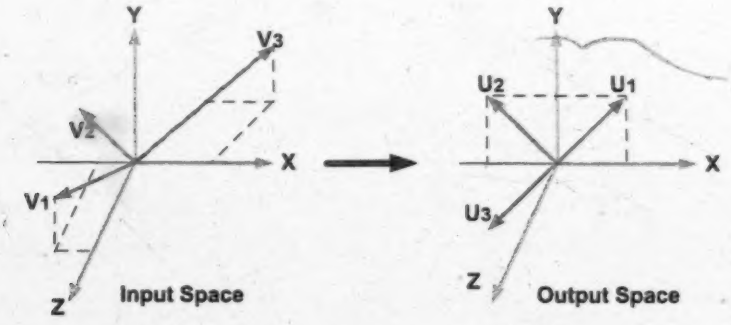

Figure 3.3: Graphical presentation of SVD decomposition of a matrix

The vectors of the $\mathbf{U}$ and $\mathbf{V}$ are shown in Figure 3.3. The first two columns of $\mathbf{U}$ matrix ($\mathbf{u}_1$ and $\mathbf{u}_2$) lie in the XY plane, whereas the third column, $\mathbf{u}_3$ points in the Z direction. Only $\mathbf{u}_1$ and $\mathbf{u}_2$ correspond to non-zero singular values, and so the range space of $\mathbf{A}$ is the XY plane.

The columns of $\mathbf{V}$, $\mathbf{v}_1$ and $\mathbf{v}_2$ span the domain of $\mathbf{A}$. If $\mathbf{A}$ is non-singular, its null space is $\mathbf{0}$ (the *trivial nullspace*). In Figure 3.3, $\mathbf{v}_1 = \begin{bmatrix} 0.4082 & 0.4082 & 0.4082 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -0.7071 & 0.7071 & 0.0000 \end{bmatrix}'$ span the domain of $\mathbf{A}$, while $\mathbf{v}_3 = \begin{bmatrix} 0.5774 & 0.5774 & -0.5774 \end{bmatrix}'$ spans the nullspace. Consequently, $\mathbf{U} \sum \mathbf{v}_3' = 0$.

3.3.1 Diagnostic Tools from SVD

As discussed in the previous section, the error matrix, $\mathbf{E}$ shows the plant-model mismatch in the sensitivity of the process outputs with respect to the inputs. If the plant-model mismatch is negligible then $\mathbf{E}$ will be a *sparse matrix*. Increasing mismatch will be reflected by the error matrix, $\mathbf{E}$ in two ways. Either the elements of the matrix will become larger, or the matrix will gradually become less sparse. These two cases are elaborated below:

- As the plant-model mismatch in one section of model increases, the plant sensitivity matrix of that section will increasingly differ from the estimated model sensitivity matrix. The substantial mismatch can be easily identified from the rest of the model by examining the error matrix $\mathbf{E}$. Processes with block diagonal sensitivity matrices are good examples for this kind of mismatch. A block-diagonal sensitivity matrix of a process indicates that the measured variables are influenced by a few manipulated variables and plant-model mismatch of one section will not have much effect on another section of the system. Scrutinizing the equations related to the erroneous block will isolate the cause of the mismatch.

- Processes with recycle streams, which transfer semi-finished product from the end of the process to the start of the processing train, are examples of integrated processes. In an integrated process, changes made to one manipulated variable ($\mathbf{x}$) can change all the dependent variables($\mathbf{y}$). The sensitivity matrix of an integrated process will be a full matrix (as opposed to a sparse matrix). The error matrix will not be a sparse matrix, like the previous case, and identifying the root cause of the mismatch will be relatively difficult. Processes with recycle streams or heat recovery systems are good examples for this category.

It should be noted that the error matrix alone gives information about the magnitude of the mismatch in the sensitivities, but does not quantify the economic impact of the mismatch in monetary units. In Equation 2.3 the $m \times n$ sensitivity matrix ($\frac{d\mathbf{y}}{d\mathbf{x}}$)

is multiplied by the $1 \times m$ profit gradient ($\frac{\partial P}{\partial y}$) to express the gradient in terms of profitability. Similarly the $m \times n$ error matrix ($\mathbf{E}$) is multiplied by the $1 \times m$ profit gradient ($\frac{\partial P}{\partial y}$) to represent the error in units of plant profitability. This collapses the error matrix ($\mathbf{E}$) into a vector component in the reduced gradient ($\nabla_r P$). This resulting vector expresses the sensitivity of the plant profit (P) to changes in the manipulated variables ($\mathbf{x}$) via the changed in the measured variables ($\mathbf{y}$). Multiplying the vector ($\frac{\partial P}{\partial y}$) and the matrix ($\frac{dy}{dx}$) takes away the opportunity to study their interaction and to find the root cause of the mismatch. It is necessary to devise techniques, which will allow the RTO user to study the interaction between them, and thereby identify the plant-model mismatches with significant economic impact.

The profit gradient ($\frac{\partial P}{\partial y}$) can be converted to a weighting factor for the error matrix ($\mathbf{E}$). The profit gradient dictates which mismatch will contribute more or less to the process performance. For example, consider one element of an error matrix, $\mathbf{E}_{11}$ to be significantly larger than a second element $\mathbf{E}_{22}$. It indicates that the equations related to $\mathbf{E}_{11}$ are misrepresenting the actual process more than the ones related to $\mathbf{E}_{22}$. This gives the wrong impression that the $\mathbf{E}_{22}$ section of the model should be given higher priority in mismatch elimination. Added to this assume that, the profit gradient corresponding to $\mathbf{E}_{11}$ is close to zero, and the profit gradient corresponding to $\mathbf{E}_{22}$ is quite large. In that case, the mismatch in $\mathbf{E}_{11}$ will cause less loss of profit than $\mathbf{E}_{22}$. Even though $\mathbf{E}_{22}$ is small, its influence on the profitability gives it higher priority in mismatch identification.

Consider a process where the profit gradient is a vector of ones, i.e., all the process outputs have equal influence on the profit function. In that case, the largest mismatch in the error matrix ($\mathbf{E}$) will be the largest contributor to profit loss. This creates the opportunity to apply algebraic tools on the error matrix ($\mathbf{E}$) for further analysis, without having to consider the profit gradient. This advantage of having the profit gradient ($\frac{\partial P}{\partial y}$) as a vector of ones ($\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$) can be achieved by re-scaling the output measurements ($\mathbf{y}$) into a new set of variables ($\mathbf{z}$) as follows. Use the profit gradient information to scale the output variable $\mathbf{y}$:

$$z_i = \frac{\partial P}{\partial y_i} y_i \tag{3.18}$$

and

$$\mathbf{z} = diag\left(\frac{\partial P}{\partial \mathbf{y}}\right) \mathbf{y} \tag{3.19}$$

The sensitivity matrix for the scaled outputs ($\mathbf{z}$) with respect to the inputs ($\mathbf{x}$) will be:

$$\frac{d\mathbf{z}}{d\mathbf{x}} = diag\left(\frac{\partial P}{\partial \mathbf{y}}\right) \frac{d\mathbf{y}}{d\mathbf{x}} \tag{3.20}$$

where $diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)$ is used to create a diagonal matrix where elements are from $\frac{\partial P}{\partial \mathbf{y}}$. Consequently, to define the mismatch matrix $\mathbf{E}$ in scaled variables, $\mathbf{z}$, the error matrix $\mathbf{E}$ is multiplied by the scaling matrix $diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)$, giving the new *profit weighted error matrix*, $\hat{\mathbf{E}}$:

$$\begin{aligned} \hat{\mathbf{E}} &= \left(\left.\frac{d\mathbf{z}}{d\mathbf{x}}\right|_{plant} - \left.\frac{d\mathbf{z}}{d\mathbf{x}}\right|_{model}\right) \\ &= diag\left(\frac{\partial P}{\partial \mathbf{y}}\right) \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} - diag\left(\frac{\partial P}{\partial \mathbf{y}}\right) \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model} \\ &= diag\left(\frac{\partial P}{\partial \mathbf{y}}\right) \left(\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} - \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model}\right) \end{aligned} \tag{3.21}$$

Note that since $\mathbf{z} = diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)\mathbf{y}$ then $\frac{dP}{d\mathbf{z}} = \frac{dP}{d\mathbf{y}}\left(diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)\right)^{-1}$. Consequently, $\frac{dP}{d\mathbf{z}} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$ as expected, and the equation for lost profit becomes:

$$\begin{aligned} \Delta P = \frac{\partial P}{\partial \mathbf{y}}.\mathbf{E} &= \frac{\partial P}{\partial \mathbf{y}} \left\{diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)\right\}^{-1} \left\{diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)\right\} \mathbf{E} \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \left[\left\{diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)\right\} \mathbf{E}\right] \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \hat{\mathbf{E}} \end{aligned} \tag{3.22}$$

It is evident from Equation 3.22 that the impact of mismatch on profitability is given by the profit weighted error matrix, $\hat{\mathbf{E}}$. Large elements in $\hat{\mathbf{E}}$ indicate large

plant-model mismatch in that section of the process contributing to large profit loss. Correlation of these large mismatches with the input-output variables can be studied by investigating the range space and null space of $\mathbf{z}$ using singular value decomposition (SVD). Two techniques of analyzing plant-model mismatch using SVD are elaborated below:

1. *Null Space of the error matrix*, $\mathbf{E}$: SVD decomposition of $\mathbf{E}$ can provide useful insight about the range space and the null space of the plant-model mismatch. The columns of $\mathbf{V}$ corresponding to the zero singular values ($\sigma_i = 0$), will be the basis for the null space of $\mathbf{E}$. As the model is improved, by eliminating plant-model mismatch, the error matrix will become increasingly sparse. A sparse matrix may have more null vectors, corresponding to zero singular values. These null vectors will identify the region of the model that is not causing any profit loss, and therefore need not be checked for model discrepancies.

2. *Decomposition of the Profit Weighted Error Matrix* $\hat{\mathbf{E}}$: SVD can also be used to decompose the profit weighted error matrix ($\hat{\mathbf{E}}$) into input-output matrices based on the range space and the domain they occupy. It can be noted from Equation 3.14 that each of the singular values are related to one range space and one domain. Multiplying the product matrices of the range space vectors ($\mathbf{u}_i$) and the domain vectors ($\mathbf{v}_i'$) by their corresponding singular values (σ_{ii}) produce a set of matrices ($\mathbf{u}_i\mathbf{v}_i'\sigma_{ii}$), where each of the matrices span one range space and one domain space. The sum of these matrices give the original profit weighted error matrix $\hat{\mathbf{E}}$, but individually these matrices can give useful information about the interaction of the mismatches in $\hat{\mathbf{E}}$.

$$\hat{\mathbf{E}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}' = \mathbf{u}_1\mathbf{v}_1'\sigma_{11} + \mathbf{u}_2\mathbf{v}_2'\sigma_{22} + \ldots\ldots + \mathbf{u}_r\mathbf{v}_r'\sigma_{rr} \qquad (3.23)$$

In Equation 3.23, $\mathbf{u}_i$ and $\mathbf{v}_i$ are the column vectors of the $\mathbf{U}$ and $\mathbf{V}$ matrix, and σ_is are the singular values. Each of these new matrices, $\mathbf{u}_r\mathbf{v}_r'\sigma_r$ span one range component and one domain component. This technique breaks down the

original profit weighted error matrix into sub-matrices, where the plant-model mismatch is clustered according to their new orientation in vector spaces. The mismatches related to an input or an output will group themselves in one column of the **U** or the **V** matrix respectively. Unless the process is highly integrated, all of the mismatch related to one fault should show up in one sub-matrix ($\mathbf{u}_r\mathbf{v}_r'\sigma_r$). The sub-matrix, corresponding to the highest singular value, contributes to the maximum profit loss, and location of the larger elements in the sub-matrix points to the part of the process model that might be contributing to the profit loss.

3.3.2 Illustrative Example Revisited

The case studies from the modified QR decomposition section are analyzed below using the SVD decomposition technique:

CASE I

Following the procedure of SVD decomposition technique, the profit weighted error matrix $\hat{\mathbf{E}}$ is calculated from the error matrix **E** using the profit gradient as its weighting.

$$\hat{\mathbf{E}} = \begin{bmatrix} 0.0112 & 0 & 0 \\ 0.2748 & 0.4717 & 0.0092 \\ -2.0378 & 0 & -0.1254 \end{bmatrix}$$

Singular value decomposition of $\hat{\mathbf{E}}$ produces the following **U**, $\mathbf{\Sigma}$, and **V** matrices.

$$\mathbf{U} = \begin{bmatrix} -0.0054 & 0.0007 & 1.0000 \\ -0.1407 & -0.9901 & -0.0000 \\ 0.9900 & -0.1406 & 0.0055 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 2.0611 & 0 & 0 \\ 0 & 0.4674 & 0 \\ 0 & 0 & 0.0007 \end{bmatrix}$$

$$\mathbf{V}' = \begin{bmatrix} 0.9976 & -0.0311 & -0.0614 \\ 0.0322 & 0.9993 & 0.0164 \\ 0.0609 & -0.0183 & 0.9980 \end{bmatrix}$$

The first column of **U** and **V**, which correspond to the largest singular value ($\sigma_{11} = 2.0611$), indicate that the largest mismatch is related to the third measured variable (product composition, x_D) and the first manipulated variable (feed flowrate, F). It shows that an error in the equations relating the final product composition and the feed flowrate is contributing to the largest profit loss. This mismatch can be linked with the inaccurate estimate of the reaction rate constant.

In the $\mathbf{\Sigma}$ matrix the first two singular values are significantly larger than the third one. The third singular value can be discarded as being the result of process noise. The remaining two singular values indicate that the process might have two separate sources of mismatch. The second column of the **U** and **V** matrix, which correspond to the second singular value ($\Sigma_{22} = 0.4674$), identifies another mismatch relating the second measured variable (product flowrate, P) with the second manipulated variable (cooling water flowrate, W). This mismatch is indicating that the heat of reaction may have been incorrectly estimated.

Equation 3.23 can be used to derive the same conclusions given above. Equation 3.23 is used to create two sub-matrices of the original profit weighted error matrix ($\hat{\mathbf{E}}$). This technique should cluster all the elements of the sensitivity matrix related to a mismatch into one of the two sub-matrices.

$$\hat{\mathbf{E}} \approx \mathbf{u}_1\mathbf{v}_1'\sigma_{11} + \mathbf{u}_2\mathbf{v}_2'\sigma_{22}$$

$$= \begin{bmatrix} 0.0112 & 0.0004 & 0.0007 \\ 0.2892 & 0.0093 & 0.0093 \\ -2.0357 & -0.0657 & -0.1242 \end{bmatrix} + \begin{bmatrix} 0.0000 & -0.0003 & 0.0000 \\ -0.0144 & 0.4624 & -0.0085 \\ -0.0020 & 0.0657 & -0.0012 \end{bmatrix}$$

In the first matrix the loss of profit is concentrated around the input-output pair of feed flowrate (F) and product composition (x_D) . The second matrix is showing another correlation between the cooling water flow-rate (W) and product flowrate (P). Both these conclusions can be explained from the physical characteristics of the process as explained earlier. Increased steam flow-rate will increase the vapor flow-rate in the column and result in as increase in the reflux rate, which in turn will increase the purity of the product (x_D).

In this case study problem the SVD based technique successfully identifies the two sources of plant-model mismatch. It can be concluded that, compared to the modified QR decomposition technique SVD based technique can shrink the search region further to find the cause of the plant-model mismatch.

CASE II

For Case II, singular value decomposition of $\hat{\mathbf{E}}$ gives the following $\mathbf{U}$, $\mathbf{\Sigma}$ and $\mathbf{V}$ matrices:

$$\mathbf{U} = \begin{bmatrix} -0.0052 & -0.0101 & 0.9999 \\ 0.1202 & -0.9927 & -0.0094 \\ 0.9927 & 0.1201 & 0.0064 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 0.2876 & 0 & 0 \\ 0 & 0.0879 & 0 \\ 0 & 0 & 0.0003 \end{bmatrix}$$

$$\mathbf{V}' = \begin{bmatrix} -0.8815 & -0.4540 & -0.1295 \\ 0.0292 & 0.2215 & -0.9747 \\ -0.4712 & 0.8630 & 0.1820 \end{bmatrix}$$

Following Equation 3.23, the first columns of $\mathbf{U}$ and $\mathbf{V}'$ matrix gives:

$$u_1 v_1' \sigma_{11} = \begin{bmatrix} 0.0013 & 0 & 0.0007 \\ -0.0305 & 0.0010 & -0.0163 \\ -0.2516 & 0.0083 & -0.1345 \end{bmatrix}$$

It clearly shows that the three elements in the bottom row are heavily weighted. The bottom row of the matrix corresponds to the composition of the product stream (x_D). This gives an indication that the mismatch is in the distillation column.

QR decomposition failed to identify the mismatch successfully. This is because, QR decomposition looks at the mismatches associated with each input one at a time. It fails to find the interaction present between them. On the other hand, SVD tries to find the correlation between the output variables as well as the input variables. The range space of the profit weighted error matrix helps in locating any underlying relation between the column vectors.

3.4 An Integrated Diagnosis Approach

The steps associated with the diagnosis techniques discussed in this chapter are further clarified followed by a simple example.

3.4.1 Steps Involved in Mismatch Diagnosis

1. Obtain the sensitivity matrices for the plant ($\frac{dy}{dx}\big|_{plant}$) and the model ($\frac{dy}{dx}\big|_{model}$), as well as the profit gradient ($\frac{\partial P}{\partial y}$).

2. Calculate the error matrix E by subtracting the plant sensitivity matrix from the model sensitivity matrix.

3. Calculate the rank of the error matrix, **E**. A full rank sensitivity matrix will indicate that all the manipulated variables are contributing to the mismatch. On the other hand, a less than full rank error matrix will indicate that there is no mismatch in some of the input-output relationships.

4. Construct a diagonal matrix of the profit gradient $\left(diag\left(\frac{\partial P}{\partial y}\right)\right)$, and use it to scale the error matrix **E**, yielding the profit weighted error matrix ($\hat{\mathbf{E}}$).

5. Conduct a singular value decomposition of the profit weighted error matrix, ($\hat{\mathbf{E}}$).

6. Inspect the singular values in matrix **S**. If the first singular value is significantly larger than the others, then most likely there is one mismatch component in the plant. If the first two or more of the singular values are of similar scale, then either (a) there is more than one mismatch in the process, or (b) the process is highly integrated and mismatch in one section of the process is contributing to the mismatch in other sections (this can happen in processes with recycle product or energy streams).

7. Locate the largest element, $\mathbf{u}_{1max}$ in the first column of the U matrix. If $\mathbf{u}_{1max}$ is significantly larger than the other elements of the column, then the mismatch is related to the manipulated variable corresponding to $\mathbf{U}_{1max}$. In some cases, more than one element of the first column of **U** matrix might be of larger than the other elements. In that case, the error is most likely related to the manipulated variable corresponding to those elements.

   Based on the singular values (i.e., larger singular value), the other columns of the **U** matrix should be chosen for analysis.

8. Similarly, the **V** matrix indicates the manipulated variable contributing to maximum profit loss. The largest element in the first column of the **V** matrix corresponds to the manipulated variable related to the mismatch.

   Combining step 7 and 8 for each of the singular values can help identify the exact elements in the sensitivity matrix with the largest mismatch.

9. Create the augmented matrix for unscaled QR decomposition by augmenting the profit gradient to the left side of the error matrix.

10. Compute the QR decomposition of the augmented matrix generating the **Q** and the **R** matrix. The elements in the first row in **R**, r_{1i} (where, n is the number of columns in the error matrix, and $i = 2, 3, .., n+1$) represent the projection

of the mismatch columns on the profit gradient. The columns of error matrix, E which are well-aligned with the profit gradient will have large values in r_{1i}, and the columns which are orthogonal will have values of r_{1i} close to zero. The column of the error matrix corresponding to the largest absolute value of r_{1i}, is the segment of the model causing maximum profit loss, and in particular identifies the input associated with the loss. Significant outputs correspond to large element in the normalized profit gradient.

3.4.2 Example

The proposed techniques are applied to a heat-exchanger network model with its RTO system. The objective for the RTO system is to maximize heat recovery from the heat-exchanger network (see Figure 3.4). Three heat-exchangers (**A**, **B** and **C**) are used to heat stream *1* by recovering waste heat from three hot streams *11*, *13* and *15*. Two flow-splitter valves (V_1 and V_2) are used to split stream *1* among the exchangers. The settings of these two valves are the manipulated variables for the process. The measured variables are the temperatures of the streams *6*, *7* and *8*.

Change in the setting of valve V_1 determines the split ratio of stream *1* between stream *2* and stream *3*. Similarly, change to the setting of valve V_2 determines the split ratio of stream *3* between stream *4* and stream *5*. Therefore, split ratio of V_1 influences the exit temperature of stream *6*, *7*, and *8*, whereas split ratio of V_2 affects only stream *7* and *8*.

The mismatch between the plant and the model is created by adding bias to the heat transfer coefficients of heat exchanger B and C. In heat-exchanger B the overall heat-transfer coefficients is decreased by 5% and in heat-exchanger C the overall heat-transfer coefficient is increased by 10%.

Since there are 2 inputs and 3 outputs the sensitivity matrix of the process will be a 3×2 matrix. The first and the second column represents the sensitivity of

the output temperature of stream *6*, *7* and *8* to the split ratio of valve V_1 and V_2 respectively. Valve V_2 has no physical influence on the temperature of stream *2*, and therefore the $(1,2)$ element of the matrix is zero.

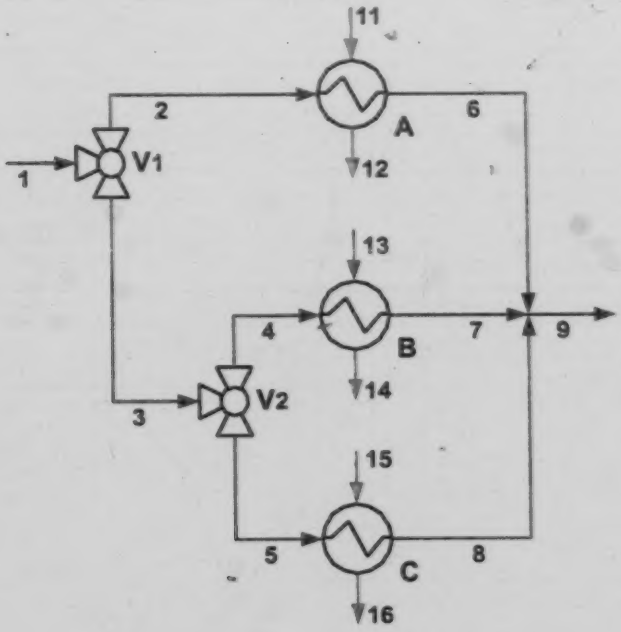

Figure 3.4: Illustrative Example: Heat Exchanger Network

The sensitivity matrix of the process, estimated from the plant operating data, as given below:

$$\left.\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right|_{plant} = \begin{bmatrix} 108 & 0.3 \\ 130 & 60 \\ 120 & 64 \end{bmatrix} \tag{3.24}$$

It should be noted that the $(1,2)$ element in the plant sensitivity matrix is not zero as expected. Noisy process data and inaccurate parameter estimation contributes to

this error.

Under the same operating conditions, the sensitivity matrix predicted by the model is,

$$\left.\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right|_{model} = \begin{bmatrix} 105 & 0 \\ 129 & 57 \\ 113 & 65 \end{bmatrix} \tag{3.25}$$

The objective of the optimizer is to maximize the temperature of stream *1*. The profit function is calculated from the heating (stream *1*) and cooling (stream *11*, *13* and *15*) cost of the process streams. The profit gradient with respect to the measured variables is,

$$\frac{\partial P}{\partial \mathbf{y}} = \begin{bmatrix} 12 & 25 & 16 \end{bmatrix} \tag{3.26}$$

The error matrix ($E = \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} - \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model}$) is calculated to be as follows:

$$E = \begin{bmatrix} 3.0 & 0.3 \\ 1.0 & 3.0 \\ 7.0 & -1.0 \end{bmatrix} \tag{3.27}$$

It is evident from Equation 3.27 that maximum absolute mismatch is in the third row first column element of the error matrix. Following the proposed methods of scaling with the profit gradient and SVD decomposition of **E** gives,

$$U = \begin{bmatrix} -0.3867 & -0.1553 & -0.9090 \\ -0.1051 & -0.9719 & 0.2107 \\ -0.9162 & 0.1771 & 0.3595 \end{bmatrix} \tag{3.28}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 7.6939 & 0 \\ 0 & 3.1455 \\ 0 & 0 \end{bmatrix} \tag{3.29}$$

$$V' = \begin{bmatrix} -0.9980 & -0.0630 \\ 0.0630 & -0.9980 \end{bmatrix} \tag{3.30}$$

Of the two singular values, the first one (σ_1) is more than twice of the second one (σ_2), but they are of the same order of magnitude. Studying the first column of $\mathbf{U}$ and $\mathbf{V}'$ (i.e., $\mathbf{U}_1$ and $\mathbf{V}'_1$ respectively), which correspond to the first singular value (σ_1), should give an understanding of the range and domain of the most significant mismatch. In vector $\mathbf{U}_1$, the third element is significantly larger than the other two. In vector $\mathbf{V}'_1$ the first parameter is much larger than the second one. This confirms the finding from the error matrix ($\mathbf{E}$) that the maximum mismatch is related to the first valve (manipulated variable) and the temperature reading of the third heat exchanger (steam 5 or stream 8). Given this relatively small model, the mismatch could be easily identified. For a larger model, the source of mismatch has to be identified from the SVD analysis.

In the preceding paragraph, SVD has been applied to the unscaled error matrix to identify the source of maximum mismatch. The largest mismatch might not be the main source of profit loss for the process. It can be noted from the profit gradient that, the influence on the second row of the error matrix on total profitability is relatively higher that the other two (i.e., $25 > 12, 16$). Therefore, it is not possible to predict whether the largest mismatch is the cause of the largest loss, just through SVD decomposition of the error matrix before . The proposed modified method of SVD decomposition and QR decomposition are elaborated below:

Diagnosis using SVD:

1. The SVD of the error matrix, $\mathbf{E}$ shows that the system does not have a non trivial null space. This indicates that mismatch is present throughout the model and both of the manipulated variables are contributing to the plant-model mismatch. If the model was close to the plant in terms of its predictions, the error matrix would have been sparse and made it easier to isolate the mismatch.

2. The diagonal form of the profit gradient and its inverse are calculated as follows,

$$diag\left(\frac{\partial P}{d\mathbf{y}}\right) = \begin{bmatrix} 12 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 16 \end{bmatrix} \tag{3.31}$$

$$diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)^{-1} = \begin{bmatrix} 0.0833 & 0 & 0 \\ 0 & 0.0400 & 0 \\ 0 & 0 & 0.0625 \end{bmatrix} \tag{3.32}$$

The product of the profit gradient and its diagonal inverse gives a vector of ones.

$$\frac{\partial P}{\partial \mathbf{y}}\left\{diag\left(\frac{\partial P}{\partial \mathbf{y}}\right)^{-1}\right\} = \begin{bmatrix} 12 & 25 & 16 \end{bmatrix}\begin{bmatrix} 0.0833 & 0 & 0 \\ 0 & 0.0400 & 0 \\ 0 & 0 & 0.0625 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \tag{3.33}$$

This ensures that the profit gradient of the new profit weighted error matrix, $\hat{\mathbf{E}}$ will be equally weighted for each of the columns of $\hat{\mathbf{E}}$.

The singular value decomposition of $\hat{\mathbf{E}}$ will be,

$$[\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \text{SVD}\left[diag\left(\frac{\partial P}{\partial \mathbf{y}}\right) \times \left(\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} - \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model}\right)\right] \tag{3.34}$$

$$\mathbf{U} = \begin{bmatrix} -0.2999 & -0.0353 & -0.9533 \\ -0.2232 & -0.9690 & 0.1061 \\ -0.9275 & 0.2446 & 0.2827 \end{bmatrix} \tag{3.35}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 120.2924 & 0 \\ 0 & 76.7378 \\ 0 & 0 \end{bmatrix} \tag{3.36}$$

$$\mathbf{V}' = \begin{bmatrix} -0.9997 & 0.0248 \\ -0.0248 & -0.9997 \end{bmatrix} \tag{3.37}$$

Analysis of the $\mathbf{\Sigma}$ matrix shows that, the singular values are again of the same order of magnitude, even though the first one (σ_1) is about one and a half times larger than the second one (σ_1). Therefore the mismatch is not isolated in a specific part of the model.

Using Equation 3.23, the profit weighted error matrix ($\hat{\mathbf{E}}$) can be decomposed into two matrices as shown in the following Equation 3.38:

$$\begin{aligned}\hat{\mathbf{E}} &= \sum_{i=1}^{n} \mathbf{u}_i \mathbf{\Sigma}_{ii} \mathbf{v}_i' \\ &= [\mathbf{u}_1 \sigma_{11} \mathbf{v}_1'] + [\mathbf{u}_2 \sigma_{22} \mathbf{v}_2'] \\ &= \begin{bmatrix} 36.1 & 0.9 \\ 26.8 & 0.7 \\ 111.5 & 2.8 \end{bmatrix} + \begin{bmatrix} -0.1 & 2.7 \\ -1.8 & 74.3 \\ 0.5 & -18.8 \end{bmatrix}\end{aligned} \tag{3.38}$$

The first column of the $\mathbf{U}$ matrix shows that the third element, which corresponds to the third measured variable (temperature of stream 8), is much larger than the other two. This indicates that the maximum loss in profit is contributed by errors related to the third measured variable. In the first column of $\mathbf{V}'$, the first element is close to one, and this means that almost all the profit loss is associated with the first input (setpoint of valve $\mathbf{V}_1$).

Keeping in mind that the two singular values are of the same scale, it is necessary to consider the second column of $\mathbf{U}$ matrix as well. The second column of $\mathbf{U}$ has the largest contribution associated with the second measured variable. In the second column of $\mathbf{V}'$, the second element is close to one. This means the second prominent mismatch is contributed by the sensitivity matrix relating the second measured variable (temperature of stream 7) and the second manipulated variable (setpoint of valve $\mathbf{V}_2$).

It is evident from the discussion above that decomposition of $\hat{\mathbf{E}}$ identified two significant mismatch components: (a) $u_1 \rightarrow y_3$, and (b) $u_2 \rightarrow y_2$. It shows

that the mismatch has a definite pattern and can be identified as pairs of input-output variables. This interaction between the input-output variables could not be determined from the error matrix and profit gradient alone.

In conclusion it can be said that the profit loss is not contributed by any definite part of the model, but it is strongly related to the third and also the second measured variable.

Diagnosis using Modified QR Decomposition:

The **Q** and the **R** matrix for the system are obtained by decomposing the modified error matrix, $\tilde{\mathbf{E}} = \left[\left(\frac{\partial P}{\partial y}\right)' \ \vdots \ \left(\frac{\partial P}{\partial y}\Big|_{plant} - \frac{\partial P}{\partial y}\Big|_{model}\right) \right]$

$$\mathbf{Q} = \begin{bmatrix} -0.3748 & 0.1785 & -0.9097 \\ -0.7809 & -0.5898 & 0.2060 \\ -0.4998 & 0.7876 & 0.3605 \end{bmatrix} \tag{3.39}$$

$$\mathbf{R} = \begin{bmatrix} -32.0156 & -5.4036 & -1.9553 \\ 0 & 5.4590 & -2.5033 \\ 0 & 0 & -0.0154 \end{bmatrix} \tag{3.40}$$

In the **R** matrix, the first terms of the second and the third column (-5.40 and -1.96) indicate how much the two columns of the error matrix, **E**, are aligned with the direction of the profit gradient. Like in the SVD method, both the terms of are similar magnitude, even though one is more than double of the other. It indicates that the plant-model mismatch is associated with both the manipulated variables with larger emphasis on the first one.

3.5 Summary

The methods developed from singular value decomposition and QR decomposition are primarily different methods of looking at the plant-model mismatch. In a large

system, it can be difficult to visualize the interaction of the process inputs and outputs, and to get a meaningful understanding of the calculated sensitivity. The techniques developed here help to redefine the complex system into more easily understandable profitability terms. It also breaks down a larger system into smaller segments based on their contribution in the profit loss. This will help the control engineer to isolate the troublesome part of the system from the overall structure and to solve it in an easily manageable size.

Chapter 4

Case Study: Exchanger Network

Chapter 2 introduced a structured approach for the detection of plant-model mismatch in a RTO system. It was extended to the diagnosis of plant-model mismatch in Chapter 3 where techniques were proposed to identify the model equations that cause the mismatch. Simple process models with RTO system were used in Chapter 2 and 3 to demonstrate the application of the proposed tools. It has offered an overview of the proposed detection and diagnosis tools from the application point of view but was not sufficient for understanding the credibilities and limitations of the methods. Therefore, it is necessary to study a more complex multivariate problem to demonstrate the efficiency of these techniques.

This case study problem has been selected based on two criteria, (a) it has to be a multi-variable nonlinear process with a model simpler than the actual process, and (b) it has to be simple enough for easy interpretation of the detection and diagnosis results. A heat-exchanger network with six heat-exchangers was chosen as the case study problem. In recent times, heat-exchanger networks have become more prevalent in process industries with increasing demand for heat integration to reduce energy consumption and to cut down process emissions. Effective application of RTO systems can boost the performance of these heat-exchanger networks quite significantly. The interaction between the measured and manipulated variables could be easily modified by changing the flow streams and the location of the split valves. This provided the flexibility to create heat-exchanger network models with different structure of the gain matrix (e.g., block diagonal matrix) and different objective functions. The

heat-exchanger models were found to be effective in studying the performance of the detection and diagnosis tools under different process configuration and operating conditions.

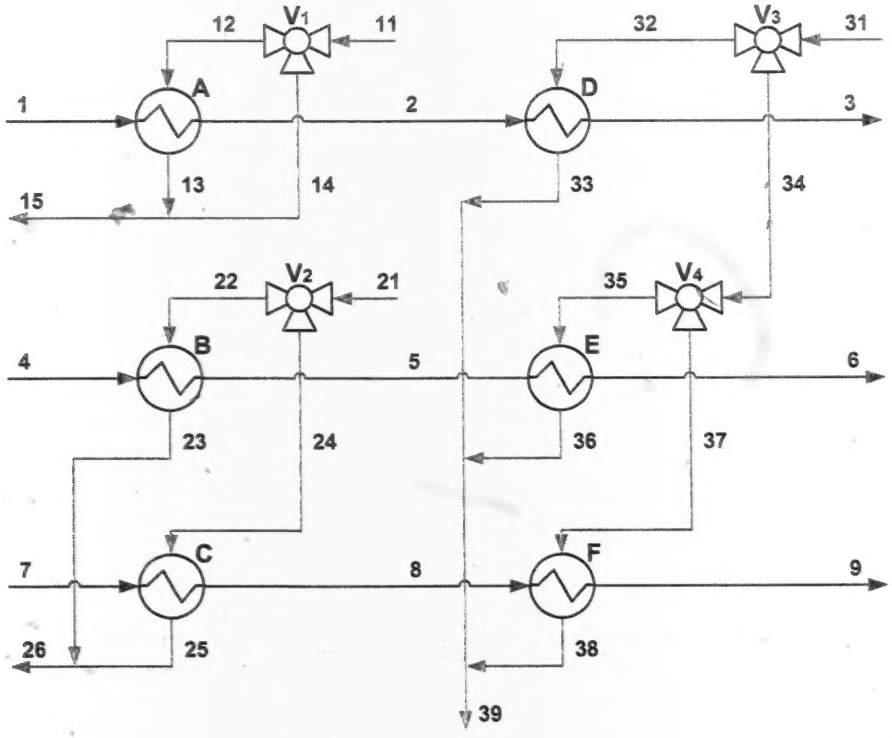

Figure 4.1: Heat Exchanger Network

Figure 4.1 presents a schematic diagram of the heat-exchanger network with stream numbers shown beside the streams. A detailed description of the process is given in Appendix B. The system consists of six heat-exchangers which are used to raise the temperatures of three cold steams *1*, *4* and *7* by exchanging heat with three hot streams *11*, *21* and *31*. The cold streams (stream *1*, *4* and *7*) leave the process as *3*, *6* and *9* and the hot streams (stream *11*, *21* and *31*) leave the process

as stream *15*, *26* and *39*. Each of these six streams has an optimum exit temperature ($\begin{bmatrix} T_3^* & T_6^* & T_9^* & T_{15}^* & T_{26}^* & T_{39}^* \end{bmatrix} = \begin{bmatrix} 295 & 286 & 290 & 294 & 264 & 300 \end{bmatrix}$), which is desired to save heating and cooling cost. The split ratio of the four valves, V_1, V_2, V_3, and V_4 are the manipulated variables of the process, which are used to control the flowrates of the hot streams through different heat exchangers. The objective of the RTO system is to ensure optimum exit temperature of the exit streams (streams *3*, *6*, *9*, *15*, *26*, and *39*) by manipulating the four valve settings (r_1, r_2, r_3, and r_4).

The following sections describe the detailed application procedure of the detection and diagnosis tools on a heat-exchanger network system. Two different operating scenarios are studied to shows the advantages as well as the limitations of these tools.

4.1 CASE STUDY I

For this case study, it was assumed that the plant and the model equations were identical except for the model parameters, i.e., the product of the overall heat transfer coefficient and the active surface area (UA), which are considered to be constant ($(UA)_i = Constant$). Excepting $(UA)_A$, the remaining five parameters are correctly estimated by the parameter estimation block. $(UA)_A$ is assumed to be 2% more than its actual value. The objective function was formulated to ensure that the process had a convex profit surface in the feasible operating region. The feasible operating region will be any condition of the four valves between completely open and closed conditions ($0 \leqslant r_i \leqslant 1$, where $i = 1, 2, 3, 4$). Therefore, the process should converge to the same optimum point independent of the initial operating conditions. It was assumed that at some given time the setpoints of the four valves were, $\begin{bmatrix} r_1 & r_2 & r_3 & r_4 \end{bmatrix} = \begin{bmatrix} 0.2000 & 0.4062 & 0.4031 & 0.4600 \end{bmatrix}$. It was desired to perform an RTO performance evaluation at this operating point using the proposed detection and diagnosis techniques.

4.1.1 Detection of Plant-Model Mismatch

Plant experiments were conducted around this operating point to calculate the sensitivity matrix of the plant, $d\mathbf{y}/d\mathbf{x}$. On the other hand, the model equations were used to calculate the sensitivity matrix of the model, $d\mathbf{y}/d\mathbf{x}\big|_{model}$. The gradients of the objective function with respect to the process outputs ($\partial P/\partial \mathbf{y}$) were calculated from Equation B.8.

$$\frac{\delta P}{\delta \mathbf{y}} = \begin{bmatrix} 0.3330 & 0.2554 & 0.1537 & 0.7674 & 0.1839 & 0.0158 \end{bmatrix} \tag{4.1}$$

Since the objective function does not have any of the manipulated variables ($\mathbf{x}$), $\frac{\delta P}{\delta \mathbf{x}}$ will be a vector of zeros.

$$\frac{\delta P}{\delta \mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.2}$$

The sensitivity matrix of the plant ($\frac{d\mathbf{y}}{d\mathbf{x}}\big|_{plant}$) calculated from the plant experiment data and the sensitivity matrix of the model ($\frac{d\mathbf{y}}{d\mathbf{x}}\big|_{model}$) estimated from the model equations were as follows:

$$\frac{d\mathbf{y}}{d\mathbf{x}}\bigg|_{plant} = \begin{bmatrix} 30.0640 & 0 & 34.4128 & 0 \\ 0 & 15.8886 & -26.2736 & 34.0956 \\ 0 & -15.0642 & -27.7486 & -30.6747 \\ -24.9428 & 0 & 0 & 0 \\ 0 & 2.5959 & 0 & 0 \\ 16.1465 & -0.3101 & 10.6090 & 3.9142 \end{bmatrix} \tag{4.3}$$

$$\frac{d\mathbf{y}}{d\mathbf{x}}\bigg|_{model} = \begin{bmatrix} 22.2509 & 0 & 23.2147 & 0 \\ 0 & 8.5324 & -16.7562 & 21.7452 \\ 0 & -3.5382 & -15.2328 & -16.8393 \\ 19.2142 & 0 & 0 & 0 \\ 0 & 4.5799 & 0 & 0 \\ 11.9503 & 1.4089 & -0.1988 & 1.4147 \end{bmatrix} \tag{4.4}$$

Based on the information given above, it was necessary to conduct a detection and diagnosis study of the plant-model mismatch. The first step of the study was to determine the existence of any plant-model mismatch. The plant-model mismatch was quantified by the angle between the profit gradient of the plant ($\nabla_r P|_{plant}$) and the profit gradient of the model ($\nabla_r P|_{model}$). The profit gradients were calculated using Equation 4.1, 4.2, 4.3, and 4.4 are given below:

$$\nabla_r P|_{plant} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} = \begin{bmatrix} -8.8755 & 2.2162 & 0.6519 & 4.0577 \end{bmatrix} \quad (4.5)$$

$$\nabla_r P|_{model} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model} = \begin{bmatrix} 22.3439 & 2.5001 & 1.1064 & 2.9893 \end{bmatrix} \quad (4.6)$$

The mismatch angle between the plant profit gradient and the model profit gradient was calculated (using Equation 2.7) as follows:

$$\theta = \frac{180^\circ}{\pi} \times \cos^{-1}\left(\frac{\nabla_r P|_{plant} \cdot \nabla_r P|_{model}}{\left\|\nabla_r P|_{plant}\right\|_2 \left\|\nabla_r P|_{model}\right\|_2}\right) = 142.19^\circ \quad (4.7)$$

As mentioned in Chapter 2, the break off point for the plant-model mismatch angle is 90°. Any angle equal or larger than 90° indicates that the optimizer has failed to identify the direction of the true optimum and its next RTO move can contribute to loss of process performance and profitability. On the other hand, as the mismatch angle moves from 90° and toward 0°, the profitability of the process increases in the subsequent RTO steps. At the operating point under consideration, i.e., $\begin{bmatrix} 0.2000 & 0.4062 & 0.4031 & 0.4600 \end{bmatrix}$, the plant-model mismatch angle was significantly larger than the critical angle of 90°. It indicated that in the next RTO step the optimizer can move the process to some less profitable operating conditions and away from the plant optimum. It is necessary to diagnose the cause of this wrong prediction by identifying the part of the model contributing to the mismatch. In the following subsections modified QR decomposition and singular value decomposition techniques (see Chapter 3) are used to diagnose this problem.

4.1.2 Modified QR Decomposition

QR decomposition is used as a numerical tool to study the orientation of the plant-model mismatch matrix with respect to the profit gradient of the plant. The difference between the plant and the model sensitivity matrices, i.e., the *error matrix* is calculated by subtracting the two matrices. The *augmented error matrix*, $\tilde{\mathbf{E}}$, is generated by adding the profit gradient as the first column to the error matrix, $\mathbf{E}$. The error matrix ($\mathbf{E}$) and the augmented error matrix ($\tilde{\mathbf{E}}$) for the given case are shown below:

$$\mathbf{E} = \begin{bmatrix} 7.8131 & 0 & 11.1981 & 0 \\ 0 & 7.3562 & -9.5174 & 12.3504 \\ 0 & -11.5259 & -12.5158 & -13.8354 \\ -44.1571 & 0 & 0 & 0 \\ 0 & -1.9840 & 0 & 0 \\ 4.1962 & -1.7190 & 10.8078 & 2.4996 \end{bmatrix} \tag{4.8}$$

$$\tilde{\mathbf{E}} = \begin{bmatrix} 0.3330 & 7.8131 & 0 & 11.1981 & 0 \\ 0.2554 & 0 & 7.3562 & -9.5174 & 12.3504 \\ 0.1537 & 0 & -11.5259 & -12.5158 & -13.8354 \\ 0.7674 & -44.1571 & 0 & 0 & 0 \\ 0.1839 & 0 & -1.9840 & 0 & 0 \\ 0.0158 & 4.1962 & -1.7190 & 10.8078 & 2.4996 \end{bmatrix} \tag{4.9}$$

QR decomposition of the augmented error matrix ($\tilde{\mathbf{E}}$) produces the following $\mathbf{Q}$ and $\mathbf{R}$ matrices:

$$\mathbf{Q} = \begin{bmatrix} -0.3671 & -0.7039 & 0.0396 & 0.3941 & 0.4479 & 0.1106 \\ -0.2816 & -0.3337 & 0.5502 & -0.6087 & -0.3155 & 0.1912 \\ -0.1694 & -0.2007 & -0.8161 & -0.4602 & -0.0335 & 0.2282 \\ -0.8461 & 0.5176 & -0.0040 & 0.1171 & 0.0051 & 0.0498 \\ -0.2027 & -0.2402 & -0.1274 & -0.0151 & -0.2987 & -0.8919 \\ -0.0174 & -0.1651 & -0.1159 & 0.4984 & -0.7807 & 0.3180 \end{bmatrix} \tag{4.10}$$

$$\mathbf{R} = \begin{bmatrix} -0.9071 & 34.4185 & 0.3131 & 0.5011 & -1.1778 \\ 0 & -29.0494 & 0.6193 & -3.9793 & -1.7566 \\ 0 & 0 & 13.9058 & 4.1690 & 17.7968 \\ 0 & 0 & 0 & 21.3533 & 0.0954 \\ 0 & 0 & 0 & 0 & -5.3845 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.11}$$

As described in Chapter **3**, the modified QR decomposition technique maps the error matrix to a new orthogonal space defined by the **Q** matrix. The columns of the **Q** matrix are the coordinates of the new basis for the range $\tilde{\mathbf{E}}$ where the profit gradient with respect to the process outputs ($\partial P/\partial y$) is the first coordinate (see column one of equation 4.10). The columns of the **Q** matrix map the columns of the augmented error matrix ($\tilde{\mathbf{E}}$) into the new vector space. The first elements of the column vectors in R matrix (i.e.,R_{11}, R_{12}, etc.) correspond to the coordinates along the first column of the **Q** matrix, which is the normalized profit gradient with respect to the process outputs ($\partial P/\partial y$). The magnitude of these elements (R_{11}, R_{12}, ...) indicate the alignment of the columns of the augmented error matrix ($\tilde{\mathbf{E}}$) with the profit gradient. Therefore, comparing the magnitudes of the elements in the first row of the **R** matrix except the first element which corresponds to the profit gradient itself can be used to find the influence of the manipulated variables on the overall plant-model mismatch.

In this case study examination of the elements in the first row of the **R** matrix showed that R_{12} was significantly larger than the rest, i.e, R_{13}, R_{14}, and R_{15}. Therefore, it can be deduced that the first column of the error matrix is more aligned to the direction of the profit gradient ($\partial P/\partial y$)) than the other columns. Since the first column of **R** is related to the first manipulated variable, i.e., the split ratio of valve V_1, the largest contribution to the plant-model mismatch came from the equations related to valve V_1. In summary, it can be said that, modified QR decomposition is an useful plant-model mismatch diagnosis tool which can be used to identify the manipulated variable(s) related to the more costly plant-model mismatches.

4.1.3 Singular Value Decomposition

Singular Value Decomposition (SVD) is another plant-model mismatch diagnosis tool. It analyzes the error matrix from a different perspective. In this method the profit gradients with respect to the outputs ($\partial P/\partial y$) are used as weighting factors and the error matrix ($\mathbf{E}$) is converted to the *profit weighted error matrix*, $\hat{\mathbf{E}}$. The profit weighted error matrix, $\hat{\mathbf{E}}$ is quite different from the error matrix, $\mathbf{E}$. Instead of showing the mismatch between the gain matrices, the profit weighted error matrix ($\hat{\mathbf{E}}$) shows the lost profit associated with each of the element of the error matrix ($\mathbf{E}$). Therefore, instead of having different units for each of the elements (e.g., the gain matrix or the error matrix $\hat{\mathbf{E}}$), the unit for all elements of the profit weighted error matrix is the unit of the profit function. The calculated *profit weighted error matrix*, $\hat{\mathbf{E}}$ for this case study is given below:

$$\hat{\mathbf{E}} = \begin{bmatrix} 2.6017 & 0 & 3.7289 & 0 \\ 0 & 1.8790 & -2.4311 & 3.1547 \\ 0 & -1.7711 & -1.9232 & -2.1259 \\ -33.8874 & 0 & 0 & 0 \\ 0 & -0.3648 & 0 & 0 \\ 0.0663 & -0.0272 & 0.1709 & 0.0395 \end{bmatrix} \tag{4.12}$$

Singular value decomposition is used as an algebraic tool to construct the orthonormal basis for the domain and the range space of the profit weighted error matrix ($\hat{\mathbf{E}}$). It decomposes the given matrix into three matrices called $\mathbf{U}$, $\mathbf{\Sigma}$ and $\mathbf{V}$. The columns of $\mathbf{U}$ represent an orthonormal set of basis vectors that span the range space of the matrix $\hat{\mathbf{E}}$. On the other hand, the columns of $\mathbf{V}$ are an orthonormal set of basis vectors for the domain of the mapping. $\mathbf{\Sigma}$ is a diagonal matrix of the singular values for the given matrix. By convention, the singular values are arranged in a descending order in the $\mathbf{\Sigma}$ matrix. Therefore, the first column of $\mathbf{V}$ corresponding to the largest singular value points to the largest source of profit loss and in that column the row with the largest element points to the manipulated variable ($\mathbf{x}$) causing that

profit loss. Similarly, the largest element in the first column of $\mathbf{U}$ points toward the measured variable ($\mathbf{y}$) contributing to the largest loss of profit. If the first two or more singular values are of the same magnitude then all the vectors in $\mathbf{U}$ and $\mathbf{V}$ should be taken under consideration.

Singular value decomposition of $\hat{\mathbf{E}}$ produces the following $\mathbf{U}$, Σ and $\mathbf{V}$ matrices:

$$\mathbf{U} = \begin{bmatrix} -0.0775 & -0.5892 & -0.5018 & -0.4189 & -0.4565 & -0.1062 \\ 0.0006 & 0.8060 & -0.3622 & -0.3349 & -0.3234 & -0.0496 \\ 0.0005 & -0.0067 & 0.7830 & -0.3906 & -0.4808 & -0.0552 \\ 0.9970 & -0.0463 & -0.0392 & -0.0324 & -0.0351 & -0.0062 \\ -0.0000 & -0.0215 & 0.0411 & -0.7362 & 0.6724 & -0.0612 \\ -0.0020 & -0.0248 & -0.0260 & -0.1292 & -0.0506 & 0.9897 \end{bmatrix} \tag{4.13}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 33.9884 & 0 & 0 & 0 \\ 0 & 5.1082 & 0 & 0 \\ 0 & 0 & 4.2982 & 0 \\ 0 & 0 & 0 & 0.4071 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.14}$$

$$\mathbf{V} = \begin{bmatrix} -1.0000 & 0.0070 & 0.0050 & 0.0004 \\ 0.0000 & 0.3005 & -0.4843 & 0.8217 \\ -0.0086 & -0.8120 & -0.5818 & -0.0460 \\ 0.0000 & 0.5004 & -0.6534 & -0.5681 \end{bmatrix} \tag{4.15}$$

$$\mathbf{U}_1\mathbf{\Sigma}_{11}\mathbf{V}_1^T = \begin{bmatrix} 2.6341 & 0 & 0.0227 & 0 \\ -0.0204 & 0 & -0.0002 & 0 \\ -0.0170 & 0 & -0.0001 & 0 \\ -33.8864 & 0 & -0.2914 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0680 & 0 & 0.0006 & 0 \end{bmatrix} \tag{4.16}$$

In the $\mathbf{\Sigma}$ matrix the first singular value ($\mathbf{\Sigma}_{11} = 33.9884$) is much larger than the other singular values indicating that the influence of the plant-model mismatch on the objective function can be clustered to a smaller subset of the input-output variables. The first columns of the $\mathbf{U}$ and $\mathbf{V}$ matrix, i.e., $\mathbf{U}_1$ and $\mathbf{V}_1$, corresponding to σ_{11} can provide useful information about the part of the sensitivity matrix related to the performance loss. In $\mathbf{U}_1$, the fourth element is significantly larger than the other elements and in the normalized vector $\mathbf{V}_1$ the first element is close to unity. These indicate that the largest contribution of the loss is related to the first manipulated variable, i.e., the split ratio of valve V_1 and the fourth measured output, i.e., the temperature of stream *15*. In this simulation it was considered that the equations represent the model were similar to the plant equations except the model parameters. Therefore, the plant-model mismatch was related to the inaccurate estimates of these model parameters. The model parameter, related to the split ratio of valve V_1 and the temperature of stream *15*, is the product of overall heat-transfer coefficient and active surface area of exchanger A, $(UA)_A$. In summary, the proposed techniques successfully identify the part of the model contributing to the profit loss, which is related to the incorrect estimation of $(UA)_A$.

From this simulation it is evident that the plant-model mismatch angle can be a simple and useful tool for the evaluation of RTO system performance. On the other hand, the diagnosis techniques based on modified QR decomposition and SVD can be very effective in identifying the equations that contribute to the loss of process performance.

4.2 CASE STUDY II

This case study considered that the product of the overall heat transfer coefficient and the active surface area for heat exchanger D, i.e. $(UA)_D$, is a biased estimate of the true value. It was assumed that the estimated $(UA)_D$ was 5% less than the

true value of 126.50 KJ/(hr.°C). The other parameters, i.e., $(UA)_A$, $(UA)_B$, $(UA)_C$, $(UA)_E$, and $(UA)_F$, were correctly estimated.

Based on the model with the erroneous parameter, the optimizer drives the process to the following stationary operating condition:

$$\begin{bmatrix} r_1 & r_2 & r_3 & r_4 \end{bmatrix} = \begin{bmatrix} 0.1815 & 0.441 & 0.421 & 0.448 \end{bmatrix} \tag{4.17}$$

Plant experiments conducted around this operating point resulted in the following sensitivity matrix of the plant ($\left.{}^{d\mathbf{y}}/_{d\mathbf{x}}\right|_{plant}$):

$$\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{plant} = \begin{bmatrix} 31.5042 & 0.0000 & 33.5677 & 0.0000 \\ 0.0000 & 14.7567 & -26.4996 & 34.2848 \\ 0.0000 & -16.1879 & -28.9369 & -30.3269 \\ -26.7922 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 4.8079 & 0.0000 & 0.0000 \\ 17.7381 & -1.3667 & 11.8406 & 3.6251 \end{bmatrix} \tag{4.18}$$

The sensitivity matrix of the model ($\left.{}^{d\mathbf{y}}/_{d\mathbf{x}}\right|_{model}$) calculated from the model equations which is given below:

$$\left.\frac{d\mathbf{y}}{d\mathbf{x}}\right|_{model} = \begin{bmatrix} 24.4367 & 0.0000 & 21.3297 & 0.0000 \\ 0.0000 & 7.7473 & -17.1405 & 22.1786 \\ 0.0000 & -3.8989 & -15.8361 & -16.5953 \\ -18.5811 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 4.3281 & 0.0000 & 0.0000 \\ 13.1480 & 0.8771 & 0.3091 & 0.9855 \end{bmatrix} \tag{4.19}$$

The gradients of the objective function with respect to the process outputs ($\frac{\delta P}{\delta y}$) were considered to be the same for both plant and model. These gradients calculated using Equation B.8 can be presented as:

$$\frac{\delta P}{\delta \mathbf{y}} = \begin{bmatrix} 0.3730 & 0.1250 & 0.1617 & 0.5282 & 0.2560 & 0.0159 \end{bmatrix} \tag{4.20}$$

The profit gradient of the plant ($\nabla_r P|_{plant}$) and the model ($\nabla_r P|_{model}$) were calculated from Equations 4.18, 4.19, and 4.20, and are given below:

$$\nabla_r P|_{plant} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \frac{d\mathbf{y}}{d\mathbf{x}}\bigg|_{plant} = \begin{bmatrix} -2.1176 & 0.4356 & 4.7175 & -0.5619 \end{bmatrix} \tag{4.21}$$

$$\nabla_r P|_{model} = \frac{\partial P}{\partial \mathbf{x}} + \frac{\partial P}{\partial \mathbf{y}} \frac{d\mathbf{y}}{d\mathbf{x}}\bigg|_{model} = \begin{bmatrix} -0.4899 & 1.4599 & 3.2575 & 0.1038 \end{bmatrix} \tag{4.22}$$

The mismatch angle between the plant gradient and the model gradient (θ) was determined to be $104.0°C$. A mismatch angle greater than 90° indicated that the process had reached a point from where it could not go any closer to the plant optimum. Therefore, it was necessary to identify the mismatch for further improvement of the process performance.

4.2.1 Modified QR Decomposition

The *error matrix* ($\mathbf{E}$) and the *augmented error matrix* ($\tilde{\mathbf{E}}$) for this case are given below:

$$\mathbf{E} = \begin{bmatrix} 7.0675 & 0.0000 & 12.2380 & 0.0000 \\ 0.0000 & 7.0093 & -9.3591 & 12.1061 \\ 0.0000 & -12.2891 & -13.1008 & -13.7316 \\ -8.2110 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.4798 & 0.0000 & 0.0000 \\ 4.5901 & -2.2438 & 11.5315 & 2.6396 \end{bmatrix} \tag{4.23}$$

$$\tilde{\mathbf{E}} = \begin{bmatrix} 0.3730 & 7.0675 & 0.0000 & 12.2380 & 0.0000 \\ 0.1250 & 0.0000 & 7.0093 & -9.3591 & 12.1061 \\ 0.1617 & 0.0000 & -12.2891 & -13.1008 & -13.7316 \\ 0.5282 & -8.2110 & 0.0000 & 0.0000 & 0.0000 \\ 0.2560 & 0.0000 & 0.4798 & 0.0000 & 0.0000 \\ 0.0159 & 4.5901 & -2.2438 & 11.5315 & 2.6396 \end{bmatrix} \tag{4.24}$$

QR decomposition of the augmented error matrix gave the following $\mathbf{Q}$ and $\mathbf{R}$ matrices:

$$\mathbf{Q} = \begin{bmatrix} -0.5144 & 0.7119 & 0.1095 & -0.0894 & -0.3249 & -0.3209 \\ -0.1724 & 0.0335 & 0.5130 & 0.7137 & 0.4301 & -0.1084 \\ -0.2231 & 0.0434 & -0.8388 & 0.4824 & 0.0921 & -0.0602 \\ -0.7285 & -0.5693 & 0.0257 & -0.2945 & 0.1667 & -0.1732 \\ -0.3531 & 0.0686 & 0.0745 & 0.1088 & -0.1884 & 0.9043 \\ -0.0220 & 0.4017 & -0.1227 & -0.3892 & 0.7986 & 0.1842 \end{bmatrix} \tag{4.25}$$

$$\mathbf{R} = \begin{bmatrix} -0.7251 & 2.2449 & 1.4128 & -2.0136 & 0.9181 \\ 0.0000 & 11.5499 & -1.1663 & 12.4627 & 0.8706 \\ 0.0000 & 0.0000 & 14.2148 & 6.1134 & 17.4044 \\ 0.0000 & 0.0000 & 0.0000 & -18.5801 & 0.9887 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 6.0491 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{4.26}$$

As mention earlier in case study *I*, the elements of the first row of the $\mathbf{R}$ matrix, except the first element, are examined to identify the manipulated variables ($\mathbf{x}$) related to the large plant-model mismatches. Unlike case study *I*, two of the elements of the first row, i.e., R_{12} and R_{14}, were of virtually identical magnitudes indicating that the largest mismatch was related to split ratio of valve V_1 and V_3. The explanation for this incorrect prediction is given after the SVD based analysis.

4.2.2 Singular Value Decomposition

The singular value decomposition of the *profit weighted error matrix*, $\hat{\mathbf{E}}$ produces the following results:

$$\hat{\mathbf{E}} = \begin{bmatrix} 2.6363 & 0.0000 & 4.5649 & 0.0000 \\ 0.0000 & 0.8761 & -1.1698 & 1.5132 \\ 0.0000 & -1.9876 & -2.1189 & -2.2209 \\ -4.3371 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.1228 & 0.0000 & 0.0000 \\ 0.0731 & -0.0357 & 0.1837 & 0.0421 \end{bmatrix} \tag{4.27}$$

$$\mathbf{U} = \begin{bmatrix} -0.8202 & 0.1691 & -0.3677 & -0.3644 & 0.1349 & -0.1116 \\ 0.1048 & 0.0748 & 0.6345 & -0.7103 & 0.2375 & -0.1412 \\ 0.3032 & -0.6771 & -0.5026 & -0.4057 & 0.1624 & -0.0777 \\ 0.4726 & 0.7121 & -0.4572 & -0.2259 & 0.0826 & -0.0514 \\ -0.0016 & 0.0108 & 0.0199 & 0.3273 & 0.9440 & 0.0349 \\ -0.0295 & 0.0133 & -0.0150 & -0.1998 & 0.0333 & 0.9786 \end{bmatrix} \tag{4.28}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 6.2187 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 4.0112 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 3.0993 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.2947 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{4.29}$$

$$\mathbf{V}^T = \begin{bmatrix} -0.6777 & -0.0820 & -0.7260 & -0.0830 \\ -0.6586 & 0.3520 & 0.5289 & 0.4032 \\ 0.3267 & 0.5026 & -0.4383 & 0.6697 \\ 0.0154 & 0.7853 & -0.0324 & -0.6181 \end{bmatrix} \tag{4.30}$$

Examination of the first row of $\mathbf{V}$, which corresponds to the largest singular value ($\sigma_{11} = 6.2187$), shows that V_{11}^T and V_{13}^T are relatively larger than the other two elements (V_{12}^T and V_{14}^T). On the other hand, the largest element in the first column of the $\mathbf{U}$ matrix is U_{11}. These indicate that a significant plant-model mismatch is related to the first and the third manipulated variables (the split ratio of valve V_1 and V_3) and the first output (temperature of stream *3*). Assuming no prior knowledge about the cause of the mismatch was available, it can be said that the plant-model mismatch was related to heat-exchanger A and D and the magnitudes of these mismatches are almost same because V_{11}^T and V_{13}^T are of similar magnitude.

In this case study, comparison of the diagnosis results with the actual cause of the mismatch show that both diagnosis techniques could identify the part of the model that caused the mismatch even though they could not definitely point to the source of the mismatch. The mismatch was introduced in heat exchanger D only and therefore, modified QR decomposition should have pointed towards the split ratio of valve V_3. Instead, the QR decomposition method indicated that the mismatch was related to valve V_1 as well as valve V_3. On the other hand, SVD based method was successful in finding the true measured variable related to the mismatch, i.e., temperature of stream *3*, unlike QR decomposition based method which pointed to the split ratio of both valve, V_1 and V_3. This can be explained by examining the physical structure of the process. In the heat-exchanger network, the cold-side output-stream of heat-exchanger A is the cold-side input-stream of exchanger D. The cold-side output temperature (the temperature of stream *3*) of exchanger D depended on the amount of heat transferred in exchanger A and D which in turn depend on the split ratios of valve V_1 or V_3. Inaccurate estimate of the overall heat-transfer coefficient of exchanger D resulted in incorrect prediction of the temperature of stream *3*. Therefore, in the gain matrix the sensitivity of the temperature of stream *3* to changes in valve V_1 or V_2 was inaccurate. Furthermore, the magnitude of the errors related to the valve settings were found to be of equal proportion. This can be explained with the fact that the same error from exchanger D was influencing the output temperature of stream *3*.

In summary it can be concluded that the proposed diagnosis tools were successful to identify the part of the model contributing to the loss of performance but cannot pinpoint finding the exact location of the mismatch. Process knowledge is invaluable in diagnosing the mismatch and helps in interpreting the behavior of RTO systems. These techniques can be be used side-by-side with with process and modelling knowledge to achieve efficient RTO system performance.

Chapter 5

Summary, Future Work & Conclusions

Automation has become a vital part of the modern process industry. These days, the industries are striving to be a low-cost producer by operating their processes in an economically optimal fashion, and automation has become a crucial means for achieving that target. Importance of process safety and environmental impact minimization is also working as a stimulant for process automation. Among the different control and optimization strategies, model-based on-line optimization has enjoyed its popularity for its promise of efficient process operation, advantages in reduced plant experimentation, and increased speed of convergence to optimal operation. This thesis focuses on such model-based on-line optimization techniques, commonly called *Real-Time-Optimization* (RTO) or *On-line Optimization*, and proposes some practical methods of RTO system performance evaluation and fault diagnosis.

5.1 Summary

The performance of a model-based RTO system is largely dependent on the accuracy of the model. For any process operating condition, an ideal model will exactly predict the process outputs. In reality, it is not possible to obtain a perfect model for a non-linear, time-variant process. The alternative solution is to use the best available model, and to modify or update it as needed to ensure correct estimation of process outputs. Depending on the influence of the input-output variables on the profit

function, some of the plant-model mismatches are more detrimental to the overall profitability of the process than others. It is crucial to identify the plant-model mismatches contributing to larger profit or performance losses, and rectify them.

The first part of this thesis (Chapter 2) presents a method for plant-model mismatch detection. The proposed method uses *reduced profit gradient* information from the plant and the model to quantify the mismatch. The reduced profit gradient shows the direction of maximum profit change, and is a function of the measured and the manipulated variables of the process. In this thesis, it is considered that all the cost factors associated with the process variables are known and they are same for both the process and the model. The only remaining unknown in the profit gradient equation is the sensitivity matrix, and the mismatch between the profit gradients are caused by the mismatch in the sensitivity matrices of the plant and its model.

The angle between the profit gradients of the plant and the model is referred as the *plant-model mismatch angle*, and it is used as a RTO performance indicator. When the profit gradient of the model is generally in the same direction as that of the plant, the mismatch angle is close to zero. An increase in the mismatch angle indicates that the RTO system is drifting away from the direction of the plant optima. A 90° angle between the gradients indicate that in the next RTO step the process will move tangentially with the *iso-profit* line (i.e. line of constant profit), and for a small RTO step the profit will remain the same. Any angle greater than 90° indicates that the RTO is moving the process away from the optima and loosing profit in every successive RTO moves.

For processes with relatively small signal-to-noise ratio, the estimation of the profit gradient of the plant will be less precise. The variance of the process noise is propagated by the parameter estimation block to the process model and thereby to the profit gradients. This replaces the gradient vector with a gradient cone, which encloses possible values of the gradient within certain statistical bounds. For the two gradient cones, the largest mismatch angle is formed by the two farthest lines on the cones. A simple bracketing and optimization routine is developed to find the largest

mismatch angle. As before, mismatch angles close to 0° indicate satisfactory RTO performance, where as angles larger than 90° indicate that the process might drifting away from its true optima.

The diagnosis section of this thesis (Chapter 3) proposes two methods for isolating the segment of the model equations causing the plant-model mismatch. These tools compare the sensitivity matrices (from the plant and the model) to identify the input-output variables related to the mismatch. The difference of the two sensitivity matrices is defined as the *error matrix*. The error matrix is studied using algebraic tools to find the source of the mismatch. The sensitivity of the profit function to the outputs ($\frac{dP}{dy}$) is used as a weighting vector to identify the segment of the error matrix contributing to larger profit losses.

The first technique uses a modified form of QR decomposition for mismatch diagnosis. Subtracting the profit gradient equations show that, the amount of profit lost due to plant-model mismatch is the product of, (i) the sensitivity of the profit to the outputs ($\frac{dP}{dy}$) and, (ii) the error matrix. The orientation of the columns of the error matrix with respect to $\frac{dP}{dy}$ determine the amount of profit loss. The modified QR decomposition maps the original error matrix into a new space, where $\frac{dP}{dy}$ is one of the basis vectors. Mapping of a column of the error matrix along $\frac{dP}{dy}$ shows the extent of profit loss, contributed by its corresponding manipulated variable.

The second technique used for mismatch diagnosis is based on *Singular Value Decomposition* (SVD). SVD of the error matrix maps the original system into a new domain. In the new domain, the vectors corresponding to zero singular values map the sensitivity matrix to its null space, and therefore do not contribute to the plant-model mismatch. This helps identify the variables that do not have any mismatch.

The sensitivity of the profit function with respect to the measured variables ($\frac{dP}{dy}$) is used as a scaling factor to convert the original error matrix to a new matrix where the elements have the unit of the profit function. This new matrix is called the *profit weighted error matrix*. SVD of the profit weighted error matrix gives the range

space and the null space of the sensitivity matrix. The vectors in the **U** and **V** matrix create clusters of the mismatch based on their affiliation with the input and output vectors. The vectors corresponding to the largest singular value identifies the measured variables and the manipulated variables related to the largest mismatch.

The concepts of each chapter were illustrated with simple case study problems, containing few equations and variables. In Chapter 4, a larger case study was used to evaluate the performance of the mismatch detection and diagnosis tools. The heat exchanger network model demonstrated the effectiveness of the detection and diagnosis tools for large scale systems.

5.2 Future Work

Some innovative ideas have evolved in the process of this research work. Some of these ideas can strengthen the credibility of the proposed technique, and some others can extend it to other branches of process control. These concepts are briefly presented below for future studies:

- Estimating the sensitivity matrix of a multi-input multi-output process is usually quite tedious. This thesis assumes that the information can be obtained from the process with minimal perturbation using one of the methods shown in Chapter 2. There is scope for further development of these perturbation techniques for RTO systems. It might be possible to design perturbation signals that will capture the process gradient information while implementing the RTO steps.

- The detection technique proposed in this thesis compares the profit gradients (from the plant and the model) for RTO performance evaluation. These gradients show the direction of the maximum profit change. However, the direction of the RTO step is not always in the direction of maximum profitability. Depending on the optimization routine used, the RTO system can choose a different

direction for the RTO move. Comparing the direction of RTO move with the profit gradients can be used as a performance evaluation tool for the optimization routine.

- This thesis assumes that, at any operating point the sensitivity of the profit gradient to the manipulated and measured variables ($\frac{dP}{d\mathbf{x}}$ and $\frac{dP}{d\mathbf{y}}$ respectively) are exactly known. This is true for linear profit functions of the form $P = \mathbf{p}'_x\mathbf{x}+\mathbf{p}'_y\mathbf{y}$ (where, $\mathbf{p}'_x$ and $\mathbf{p}'_y$ are the profit weightings with respect to the measured and the manipulated variables). Consider a process where the market price of a product stream depends on the concentration of one component of the product. The profit gradient related to that stream is linked with two measured variables, i.e. $P = f$(Flow rate, Composition). When calculating dP/d(Flow rate) the composition will have to be considered constant, and vice versa. Since there are two concentration (one from the plant and the other from the model) for any operating point, the question will be which concentration to use, when evaluating dP/d(Flow rate). In this case, $\frac{dP}{d\mathbf{y}}\Big|_{plant} \neq \frac{dP}{d\mathbf{y}}\Big|_{model}$. For industrial systems, the relation between the profit and the variables are quite complex and they can lead to similar complicated situations. This shows that, there is scope to improve the proposed detection and diagnosis techniques by using independent profit gradients for the plant and the model.

- Similar to the RTO systems, the *Model Predictive Control* (MPC) systems use a process model for output prediction. Therefore, it is possible to generate gradient information for the MPC model and the process. The detection and diagnosis techniques proposed for RTO system can be adopted and extended for performance analysis of the MPC system.

- Depending on the operating condition of a plant, its inequality constraints can switch from active to inactive form, and vice versa. An extra active constraint will eliminate an independent variable by making it dependant on other variables. This shrinks the dimension of the reduced profit gradient (dim $(\nabla_r P)$) by one. This thesis only considered problem where dim $(\nabla_r P)$ remained constant.

This leaves a new path to explore in future studies.

5.3 Conclusion

This thesis presents several key concepts on performance evaluation of model-based optimization system. It points out that RTO systems have plant-model mismatches. Minimizing these mismatches is crucial for effective long term application of the optimization systems. The proposed gradient angle measurement technique can be implemented as an off-line system to monitor RTO system operation. It can work as a simple but effective measuring stick for the control engineer, giving vital information on changes in plant profitability.

Similarly, the off-line diagnosis techniques can provide concise and meaningful understanding of the plant-model mismatch problem without going into the detailed of the equations and coding of the total RTO system. With few plant experiments, the control engineer can identify the section of the plant causing the mismatch and can take steps to generate extra profit through RTO performance improvement.

The only shortcoming of using these techniques are the plant experiments needed for gradient estimation. Considering the benefits of plant-model mismatch minimization, the cost of plant experiments should be relatively insignificant.

Bibliography

Bamberger, W. and R. Isermann (1977). Adaptive on-line steady-state optimization of slow dynamic processes. *Automatica* **14**, 223–230.

Bates, D. M. and D. G. Watts (1988). *Nonlinear Regression Analysis & Its Applications.* John Wiley & Sons. New York, U.S.A.

Beveridge, Gordon S. G. and Robert S. Schechter (1970). *Optimization: Theory and Practice.* McGraw-Hill Book Company. The University of Texas, Texas, U.S.A.

Box, M.J. (1970). Improved parameter estimation. *Technometrics* **12**(2), 219–229.

Campbell, Hugh G. (1977). *An Introduction to Matrices, Vectors, and Linear Programming.* Prentice-Hall Inc.. Englewood Cliffs, New Jersey, U.S.A.

Cutler, C. R. and R. T. Perry (1983). Real time optimization with multivariable control is required to maximize profits. *Computers and Chemical Engineering* **7**(5), 663–667.

Darby, M.L. and D.C. White (1988). On-line optimization of complex process units. *Chemical Engineering Progress* pp. 51–59.

Fletcher, R. (1987). *Practical Methods of Optimization.* John Wiley & Sons. New York, U.S.A.

Forbes, J. F., J. F. MacGregor and T. E. Marlin (1994). Model adequacy requirement for optimizing plant operation. *Computers and Chemical Engineering* **18**(6), 497–510.

Forbes, J. Fraser (1994). Model Structure And Adjustable Parameter Selection for Oeprations Optimization. PhD thesis. McMaster University. Canada.

Fu, Li-Chen, Wei-Der Chang, Dung-Ming Chuang, Te-Son Kuo, Tze-Chien Wang and Chi-Wang Tsai (1997). Model adequacy requirement for optimizing plant operation. *American Control Conference.*

Gill, P. E., W. Murray and M. H. Wright (1981). *Practical Optimization.* Academic Press. New York, U.S.A.

Golden, Melinda P. and B. Erik Ydstie (1989). Adoptive extremum control using approximate process models. *AIChE Journal* **35**(7), 1157–1169.

Golub, Gene H. and Charles F. Van Loan (1990). *Matrix Computations Second Edition.* The John Hopkins University Press. Baltimore, U.S.A.

Holman, J. P. (1972). *Heat Transfer.* McGraw-Hill. New York, U.S.A.

Korchinski, W. J., T. W. Hoffman and D. W. Hoffman (1998). Tips for control engineers. *Hydrocarbon Processing* (September), 113–118.

Krishnan, S., G. W. Barton and J.D. Perkins (1992). Robust parameter estimation in on-liine optimization - part i. methodology and simulated case study. *Computers and Chemical Engineering* **16**(6), 545–562.

Mansour, M. and J. E. Ellis (2003). Comparison of methods for estimating real process derivatives in on-line optimization. *Applied Mathematical Modelling* **27**, 275–291.

McFarlane, R. C. and D. W. Bacon (1989). Emperical strategies for open-loop on-line optimization. *The Canadian Journal of Chemical Engineering* **67**, 665–677.

Merriam-Webster (1994). *The Merriam-Webster Dictionary.* Merriam-Webster Incorporated. Springfield, MA, U.S.A.

Noble, B. and J. W. Daniel (1977). *Applied Linear Algebra.* Prentice-Hall. Englewood Cliffs, NJ, U.S.A.

Pfaff, George (2001). Information generation for real-time optimization. Master's thesis. University of Alberta.

Press, William H., Brian P. Flannery, Saul A. Teukolsky and William T. Vetterling (1977). *Numerical Recipes in C : The Art of Scientific Computing.* Prentice-Hall. Englewood Cliffs, NJ, U.S.A.

Protter, Murray H. and Charles B. Morrey (1964). *Modern Mathematical Analysis.* Addison-Wesley (Canada) Limited. Don Mills, Ontario, Canada.

Roberts, P. D. (1979). An algorithm for steady-state system optimization and parameter estimation. *International Journal of Systems Science* **10**(7), 719–734.

Seborg, Dale E., Thomas F. Edgar and Duncan A. Mellichamp (1989). *Process Dynamics and Control.* John Wiley & Sons. New York, U.S.A.

Sourander, M.l., M. Kolari, J. C. Poje and D. C. White (1994). Control and optimization of olefin-cracking heater. *Hydrocarbon Processing* pp. 63–69.

White, D. C. (1998). Online optimization: What have we learned. *Hydrocarbon Processing* pp. 55–59.

Zhang, Yale, D. Monder and J. Fraser Forbes (2002). Real-time optimization under parametric uncertainty: A probability constrained approach. *Journal of Process Control* **12**(3), 373–389.

Appendix A

Heat Exchanger Network Model

In Chapter 2 a simple heat-exchanger network with three heat-exchanger is used to illustrate the application of the proposed plant-model mismatch diagnosis techniques. Heat-exchanger networks are extensively used in modern process industry to minimize heating and cooling cost of the process streams. To adapt with changing operating conditions, they are designed with some operating flexibilities. Usually, the flowrates of the streams can be manipulated to maximize heat recovery. Efficient operation of these systems can significantly reduce the operating cost of the process. Modern process industries are using RTO system to optimize operation of these systems to harness their benefits.

A.1 Process Description

Figure A.1 shows the heat exchanger network used in Chapter 2 with stream numbers shown beside the streams. The cold feed, stream *1*, is split into two stream (i.e., stream *2* and stream *3*) by a two-way-split valve V_1. Stream *2* and *3* pass through the cold sides of the heat-exchanger A and B respectively and merge to form stream *6*.

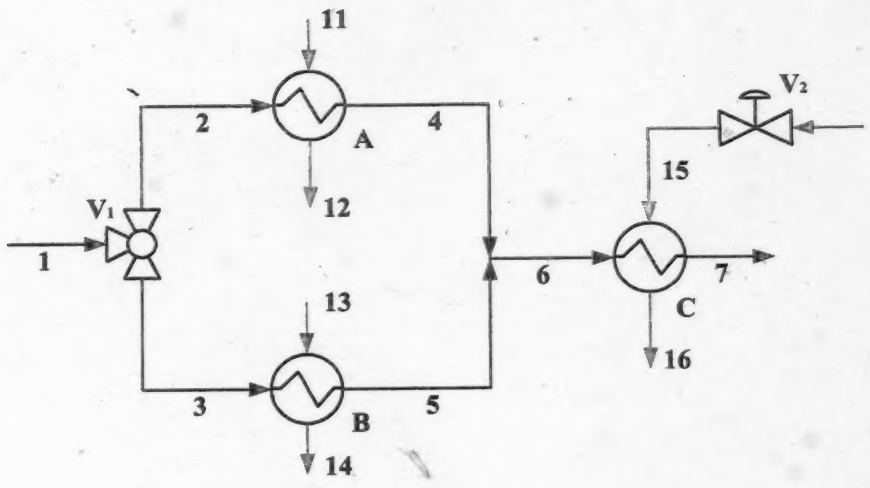

Figure A.1: Heat Exchanger Network (Numbers indicate stream number)

In heat-exchanger A, stream *11* delivers heat energy to stream *2* and leaves as stream *12*. Similarly, stream *13* loses heat to stream *3* and exits exchanger B as stream *14*. Stream *6* is further heated in heat-exchanger C before leaving the network as stream *7*. Stream *15*) serves as the heating fluid for the heat exchanger, C.

The objective of the system is to raise the temperature of the cold feed stream *1* to a higher temperature (in this case, $330°C$) by recovering maximum possible amount of heat energy from stream *11* and *13*, and supplying the rest of the heat from the heating fluid stream *15*. Two manipulated variables available are the split ratio of valve V_1 and the opening of valve V_2. The split ratio of valve V_1, r_1, is defined as the fraction of stream *1* flowing as stream *2*. The rest of the fluid is going into stream *3* (see Equation A.10). The opening of valve V_2, r_2, is scaled between zero and one. Zero indicates that the valve is completely closed while one corresponds to the completely open condition of the valve (see Equation A.24). For both valves a linear relationship between valve position and fluid flow is considered.

Heat capacity for all streams including that of process heating fluid are considered

to be unity. The flowrate and temperature of the 'i'th stream is represented as F_i and T_i, respectively. The flowrate and temperature of the streams are given in Equation A.9, A.13, A.14, A.15, A.16, A.17, A.24, and A.25.

Neglecting the resistance to heat transfer within the metal piping, the overall heat transfer coefficient can be expressed as a function of the shell side and the tube side heat transfer coefficients (Equation A.1. The heat transfer coefficient equation can be simplified to define the heat transfer rate as a function of the fluid flowrate only (Equation A.2) (Forbes, 1994; Holman, 1972):

$$U_{overall} = \left(\sum_i \frac{1}{h_i}\right)^{-1} = \left(\frac{1}{h_{\text{shell side}}} + \frac{1}{h_{\text{tube side}}}\right)^{-1} \tag{A.1}$$

$$h_i \propto F_i^{0.6} \tag{A.2}$$

In all heat-exchangers, the hot and cold fluids flow in counter current direction to each other and the driving force for the heat transfer between the two sides is expressed ~~by the log~~ *mean temperature difference* (LMTD), ΔT_{lm}. Assuming no heat is lost to the surrounding, the total amount of heat transferred (Q) can be estimated by the following equation:

$$Q = UA\Delta T_{lm} = F_i C_{p,i} \Delta T \tag{A.3}$$

where,

Q = Total amount of heat transferred

U = Overall heat transfer coefficient

A = Total heat transfer surface area

ΔT_{lm} = Log mean temperature difference between the two sides

F_i = Flowrate of the i^{th} fluid stream

$C_{p,i}$ = Heat capacity of the fluid

ΔT = Temperature change of the fluid

Based on Equation A.3, six energy balance equations (two for each heat-exchanger) are written (Equation A.18, A.19, A.21, A.22, A.26, and A.27).

A.2 Parameter Estimation and Optimization

The overall heat transfer coefficients, U_A, U_B and U_C, and active heat transfer areas A_A, A_B, and A_C for heat-exchanger A, B and C respectively, are considered to be unknown. The product of overall heat transfer coefficient and heat transfer area are represented by three adjustable parameters, β_A, β_B, and β_C using the following relationships:

$$(UA)_A = \beta_A \tag{A.4}$$

$$(UA)_B = \beta_B \tag{A.5}$$

$$(UA)_C = \beta_C \tag{A.6}$$

This is a reasonable assumption from practical point of view because in an actual process U and A are never exactly known. In the model, Equation A.4, A.5, and A.6 replace Equation A.28, A.29, and A.30, and the rest of the plant equations are considered to be identical for the model. In the parameter estimation block (see Chapter 1), the model parameters, β_A, β_B, and β_C are estimated in every RTO interval from the previous steady state process data. Considering all flow and temperature measurements are available, the parameter estimation block can be formulated as a nonlinear least squares problem as follows:

$$\min_{\beta} \left(\begin{array}{l} \left(T_4 - \hat{T}_4\right)^2 + \left(T_{12} - \hat{T}_{12}\right)^2 + \left(T_5 - \hat{T}_5\right)^2 + \\ \left(T_{14} - \hat{T}_{14}\right)^2 + \left(T_7 - \hat{T}_7\right)^2 + \left(T_{16} - \hat{T}_{16}\right)^2 \end{array} \right) \tag{A.7}$$

subject to:

The energy balance equations

(Equation A.13, A.14, A.16, A.17, A.21, A.22, A.26, A.27 and A.28)

where $\hat{T}_i$ is the predicted temperature for the stream 'i'. $\hat{T}_i$ is estimated from the model equations.

The optimizer block uses the process model to estimate the best operating condition to maximize the objective function. The optimization problem is written as:

$$\begin{aligned} &\min_{r_1, r_2} (T_7 - 330)^2 - T_6 \\ &\textit{subject to:} \\ &\mathbf{f}_m(\mathbf{x}, \mathbf{y}, \beta) = \mathbf{0} \end{aligned} \tag{A.8}$$

where $\mathbf{f}_m(\mathbf{x}, \mathbf{y}, \beta) = \mathbf{0}$ are the model equations.

A.3 Process Equations

$$F_1 = F_2 + F_3 = F_6 = F_7 = 100 \tag{A.9}$$

$$r_1 = \frac{F_2}{F_1} \tag{A.10}$$

$$F_4 = F_2 \tag{A.11}$$

$$F_5 = F_3 \tag{A.12}$$

$$F_{11} = F_{12} = 200 \tag{A.13}$$

$$F_{13} = F_{14} = 120 \tag{A.14}$$

$$T_1 = T_2 = T_3 = 250 \tag{A.15}$$

$$T_{11} = 320 \tag{A.16}$$

$$T_{13} = 341 \tag{A.17}$$

$$T_4 = T_2 + \frac{(UA)_A}{F_2} \frac{\{(T_{11} - T_4) - (T_{12} - T_2)\}}{\ln \frac{(T_{11} - T_4)}{(T_{12} - T_2)}} \tag{A.18}$$

$$T_5 = T_3 + \frac{(UA)_B}{F_3} \frac{\{(T_{13} - T_5) - (T_{14} - T_3)\}}{\ln \frac{(T_{13} - T_5)}{(T_{14} - T_3)}} \tag{A.19}$$

$$T_6 = \frac{F_4 T_4 + F_5 T_5}{F_4 + F_5} \tag{A.20}$$

$$T_{12} = T_{11} - \frac{(UA)_A}{F_{11}} \frac{\{(T_{11} - T_4) - (T_{12} - T_2)\}}{\ln \frac{(T_{11} - T_4)}{(T_{12} - T_2)}} \tag{A.21}$$

$$T_{14} = T_{13} - \frac{(UA)_B}{F_{13}} \frac{\{(T_{13} - T_5) - (T_{14} - T_3)\}}{\ln \frac{(T_{13} - T_5)}{(T_{14} - T_3)}} \tag{A.22}$$

$$r_2 = \frac{F_{15}}{250} \tag{A.23}$$

$$F_{16} = F_{15} \tag{A.24}$$

$$T_{11} = 350 \tag{A.25}$$

$$T_7 = T_6 + \frac{(UA)_C}{F_6} \frac{\{(T_{15} - T_7) - (T_{16} - T_6)\}}{\ln \frac{(T_{15} - T_7)}{(T_{16} - T_6)}} \tag{A.26}$$

$$T_{16} = T_{15} - \frac{(UA)_C}{F_{15}} \frac{\{(T_{15} - T_7) - (T_{16} - T_6)\}}{\ln \frac{(T_{15} - T_7)}{(T_{16} - T_6)}} \tag{A.27}$$

$$(UA)_A = \frac{1}{\frac{1}{11.7F_{11}^{0.6}} + \frac{1}{24.2F_2^{0.6}}} \tag{A.28}$$

$$(UA)_B = \frac{1}{\frac{1}{12.0F_{13}^{0.6}} + \frac{1}{23.2F_3^{0.6}}} \tag{A.29}$$

$$(UA)_C = \frac{1}{\frac{1}{15.0F_{15}^{0.6}} + \frac{1}{28.0F_6^{0.6}}} \tag{A.30}$$

A.4 Simulation Results

The simulation results of the process show that the RTO system quickly moves the process toward its optimum operating condition ($[0.492, 0.167]$). In the first few RTO steps, the RTO system successfully steers the process toward the process optimum. In the vicinity of the optima it fails to converge to the true optimum (see Figure A.2). Due to the plant-model mismatch caused by ignoring the flow dependence of the heat transfer coefficients and estimating the values of the coefficients directly. The RTO system considers the model optimum ($[0.504, 0.168]$) to be the process optimum.

Two major measured variables, the temperature of stream *6* and *7*, are shown in Figure A.3. From this figure it can be noted that the temperature of stream *6* increases in every successive RTO interval as the RTO system makes better estimate of r_1, the split ratio in valve V_1. The controlled variable of the process, temperature of stream *7*, was initially higher than the desired temperature of $330°C$. The RTO system brings down this temperature and level out slightly off the desired temperature, $330°C$. Figure A.4 shows the plant-model mismatch angle at each of the RTO setpoints given in Figure A.3. In the first four RTO steps, the mismatch angles are close to zero eventhough the setpoint changes in r_1 and r_2 are quite significant. On the other hand, the mismatch angle is much larger in the following steps where the change of manipulated variables are negligible. This can be explained by the fact that

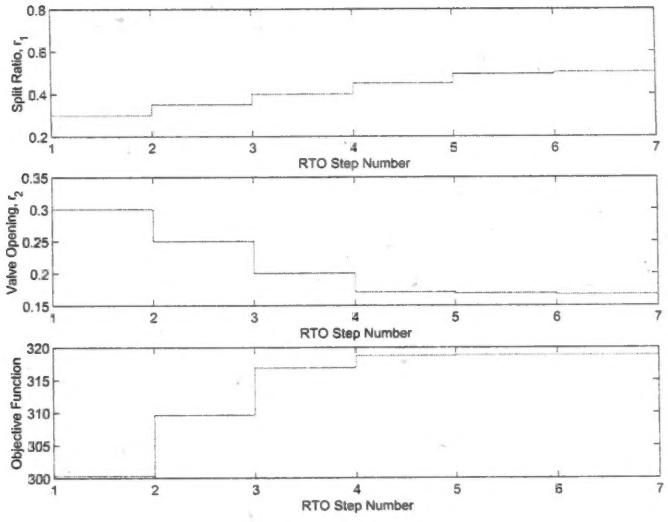

Figure A.2: RTO setpoints and their corresponding value of the profit function

when the operating point is far away from the plant and model optima, their general directions are same; whereas, when it operates close to the optima their directions can become significantly different.

Figure A.5 and A.6 present the contour plot and the three dimensional surface plot of the profit function, respectively. These plots show that the process has a local optimum in its designated operating region.

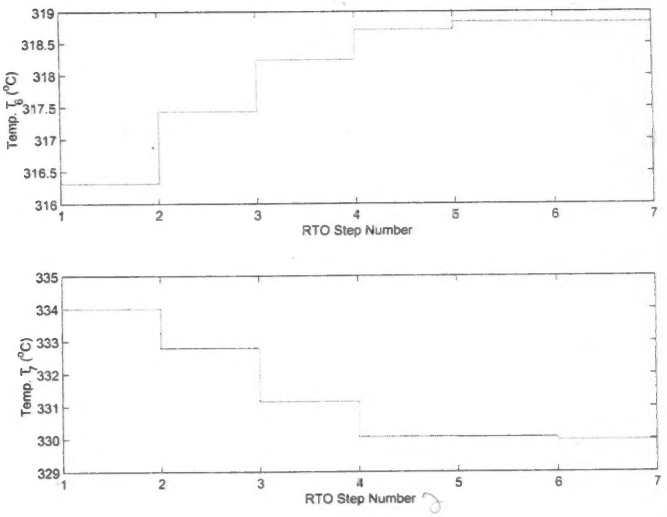

Figure A.3: The two major measured variables (T_6 and T_7)

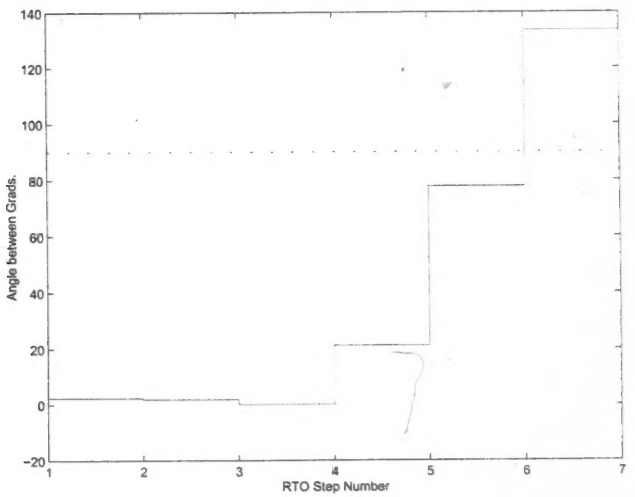

Figure A.4: The plant-model mismatch angle at different RTO setpoints

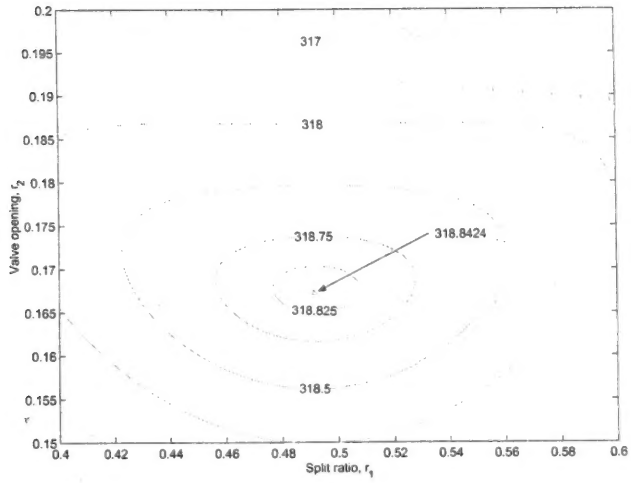

Figure A.5: Profit contour of the heat exchanger network

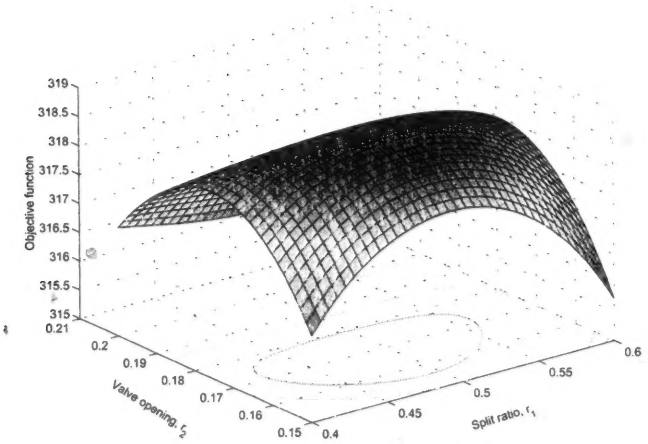

Figure A.6: 3D profit surface and contour of the heat exchanger network

Appendix B

Case Study Model

The detailed description of the six heat-exchanger network system described in Chapter **4** is provided in this appendix. This heat-exchanger network model is an enlarged form of the model presented in Chapter 2 and Appendix A. The objective for using this model is to evaluate the performance of the diagnosis technique for large systems with interaction between the process variables.

B.1 Process Description

A schematic diagram of the network is given in Figure 4.1. In this system three cold streams (stream *1*, *4* and *7*) exchange heat with three hot streams (stream *11*, *21* and *31*) to bring the temperature of the streams to their desired levels as indicated by the objective function. Four flow-splitter valves, V_1, V_2, V_3 and V_4, are used to regulate the flows of hot fluids through the exchanger network. Similar to the heat-exchanger example given in Chapter 2, the opening of the valves are scaled between zero and one, where zero is completely closed and one is completely open condition. The split ratios for the valves are r_1, r_2, r_3, and r_4, respectively (see Equation B.15, B.16, B.17

and B.18). These four split ratios are the manipulated variables of this process. The flowrate and temperature of the process streams are given in Table B.1.

Table B.1: Flowrate and temperature of the streams

Stream No.	**Flowrate (Kgs/sec.)**	**Temperature (°C)**
1	100	250.0
4	95	245.0
7	110	255
11	250	300
21	200	280
31	200	320

In this heat-exchanger network each of the cold streams passes through two heat exchangers before leaving the system (Equation B.9, B.10, and B.11). Stream *1* exchanges heat energy in exchanger A and D, and leave the system as stream *3*. The feed stream *4* is heated in exchanger B and E and it leaves the process as stream *6*. Steam *7* leaves the system as stream *9* after exchanging heat in exchanger C and F. On the hot side of the exchangers the fluid flows depend on the valve settings. A fraction of stream *11* passes through the exchanger A and merge with the rest to form stream *15* (Equation B.12). Split valve V_2 divides stream *21* between exchanger B and C. The outlet streams from the hot sides of exchanger B and C merge as stream *26* before leaving the system (Equation B.13). Stream *31* has the highest temperature and it is used to raise the temperature of all three cold streams to their final target temperatures. Valves, V_3 and V_4, are used to distribute stream *31* among

exchanger D, E and F. After delivering heat to the exchangers, the fractions of stream *31* merge again to form stream *39* (Equation B.14).

Based on Equation A.3, twelve energy balance equations are written for the six heat exchangers (see Equations B.25 to B.36). The overall heat transfer coefficients of the heat-exchanger networks are calculated using Equations A.1 and A.2 (see Equations B.40, B.41, B.42, B.43, B.44, and B.45).

B.2 Parameter Estimation and Optimization

Similar to the model given in Appendix A, the product of overall heat transfer coefficients and heat transfer areas, i.e., (UA), for all the heat exchangers are considered to be unknown to the RTO system. The RTO system considers all UAs as constant (see Equation B.1, B.2, B.3, B.4, B.5, B.6) and a parameter estimation block is used to estimate UA from the previous operating data in each RTO interval.

$$(UA)_A = \beta_A \tag{B.1}$$

$$(UA)_B = \beta_B \tag{B.2}$$

$$(UA)_C = \beta_C \tag{B.3}$$

$$(UA)_D = \beta_D \tag{B.4}$$

$$(UA)_E = \beta_E \tag{B.5}$$

$$(UA)_F = \beta_F \tag{B.6}$$

There is a desired temperature for each of the six process streams (Streams *3*, *6*, *9*, *15*, *26*, and *39*) coming out of the heat exchanger network. If the temperature

of the streams go above or below that desired temperature, it will contribute to the economic loss in terms of heating and cooling costs. The desired temperatures of the streams are as follows:

$$\begin{bmatrix} T_3^* & T_6^* & T_9^* & T_{15}^* & T_{26}^* & T_{39}^* \end{bmatrix} = \begin{bmatrix} 295 & 286 & 290 & 294 & 264 & 300 \end{bmatrix} \tag{B.7}$$

More emphasis is given on maintaining the exit temperatures of the cold streams over that of the hot stream by introducing. The weighting factor for cold streams (*3*, *6*, and *9*), and hot stream (15, 26, and 39) are set to 1, 1, 1, 0.5, 0.5 and 0.01 respectively. Therefore, the weighting vector is $\begin{bmatrix} 1 & 1 & 1 & 0.5 & 0.5 & 0.01 \end{bmatrix}$. The objective function of the optimization block is given below:

$$\min_{r_1,r_2,r_3,r_4} \left(\begin{array}{l} \left(T_3 - \hat{T}_3\right)^2 + \left(T_6 - \hat{T}_6\right)^2 + \left(T_9 - \hat{T}_9\right)^2 + \\ 0.5\left(T_{15} - \hat{T}_{15}\right)^2 + 0.5\left(T_{26} - \hat{T}_{26}\right)^2 + 0.01\left(T_{39} - \hat{T}_{39}\right)^2 \end{array} \right) \tag{B.8}$$

Subject to:

$$f(x, y, \beta) = 0$$

$$g(x, y, \beta) > 0$$

where, $\mathbf{f}(\mathbf{x}, \mathbf{y}, \beta) = \mathbf{0}$ and $\mathbf{g}(\mathbf{x}, \mathbf{y}, \beta) > \mathbf{0}$ are the equality and inequality constraints of the model.

B.3 Process Equations

$$F_1 = F_2 = F_3 = 100 \tag{B.9}$$

$$F_4 = F_5 = F_6 = 95 \tag{B.10}$$

$$F_7 = F_8 = F_9 = 110 \tag{B.11}$$

$$F_{11} = F_{12} + F_{14} = F_{13} + F_{14} = F_{15} = 250 \tag{B.12}$$

$$F_{21} = F_{22} + F_{24} = F_{23} + F_{24} = F_{23} + F_{25} = F_{26} = 200 \tag{B.13}$$

$$\begin{aligned} F_{31} &= F_{32} + F_{34} = F_{33} + F_{34} = F_{33} + F_{35} + F_{37} \\ &= F_{33} + F_{36} + F_{37} = F_{33} + F_{36} + F_{38} = F_{39} = 200 \end{aligned} \tag{B.14}$$

$$r_1 = \frac{F_{12}}{F_{11}} \tag{B.15}$$

$$r_2 = \frac{F_{22}}{F_{21}} \tag{B.16}$$

$$r_3 = \frac{F_{32}}{F_{31}} \tag{B.17}$$

$$r_4 = \frac{F_{35}}{F_{34}} \tag{B.18}$$

$$T_1 = 250 \tag{B.19}$$

$$T_4 = 245 \tag{B.20}$$

$$T_7 = 255 \tag{B.21}$$

$$T_{11} = 300 \tag{B.22}$$

$$T_{21} = 280 \tag{B.23}$$

$$T_{31} = 320 \tag{B.24}$$

$$T_2 = T_1 + \frac{(UA)_A}{F_1} \frac{\{(T_{12} - T_2) - (T_{13} - T_1)\}}{\ln \frac{(T_{12} - T_2)}{(T_{13} - T_1)}} \tag{B.25}$$

$$T_5 = T_4 + \frac{(UA)_B}{F_4} \frac{\{(T_{23} - T_4) - (T_{22} - T_5)\}}{\ln \frac{(T_{23} - T_4)}{(T_{22} - T_5)}} \tag{B.26}$$

$$T_8 = T_7 + \frac{(UA)_C}{F_7} \frac{\{(T_{24} - T_8) - (T_{25} - T_7)\}}{\ln \frac{(T_{24} - T_8)}{(T_{25} - T_7)}} \tag{B.27}$$

$$T_3 = T_2 + \frac{(UA)_D}{F_1} \frac{\{(T_{32} - T_3) - (T_{33} - T_2)\}}{\ln \frac{(T_{32} - T_3)}{(T_{33} - T_2)}} \tag{B.28}$$

$$T_6 = T_5 + \frac{(UA)_E}{F_4} \frac{\{(T_{35} - T_6) - (T_{36} - T_5)\}}{\ln \frac{(T_{35} - T_6)}{(T_{36} - T_5)}} \tag{B.29}$$

$$T_9 = T_8 + \frac{(UA)_F}{F_7} \frac{\{(T_{37} - T_9) - (T_{38} - T_8)\}}{\ln \frac{(T_{37} - T_9)}{(T_{38} - T_8)}} \tag{B.30}$$

$$T_{13} = T_{12} + \frac{(UA)_A}{F_{12}} \frac{\{(T_{12} - T_2) - (T_{13} - T_1)\}}{\ln \frac{(T_{12} - T_2)}{(T_{13} - T_1)}} \tag{B.31}$$

$$T_{23} = T_{22} + \frac{(UA)_B}{F_{22}} \frac{\{(T_{23} - T_4) - (T_{22} - T_5)\}}{\ln \frac{(T_{23} - T_4)}{(T_{22} - T_5)}} \tag{B.32}$$

$$T_{25} = T_{24} + \frac{(UA)_C}{F_{24}} \frac{\{(T_{24} - T_8) - (T_{25} - T_7)\}}{\ln \frac{(T_{24} - T_8)}{(T_{25} - T_7)}} \tag{B.33}$$

$$T_{33} = T_{32} + \frac{(UA)_D}{F_{32}} \frac{\{(T_{32} - T_3) - (T_{33} - T_2)\}}{\ln \frac{(T_{32} - T_3)}{(T_{33} - T_2)}} \tag{B.34}$$

$$T_{36} = T_{35} + \frac{(UA)_E}{F_{35}} \frac{\{(T_{35} - T_6) - (T_{36} - T_5)\}}{\ln \frac{(T_{35} - T_6)}{(T_{36} - T_5)}} \tag{B.35}$$

$$T_{38} = T_{37} + \frac{(UA)_F}{F_{37}} \frac{\{(T_{37} - T_9) - (T_{38} - T_8)\}}{\ln \frac{(T_{37} - T_9)}{(T_{38} - T_8)}} \tag{B.36}$$

$$T_{15} = \frac{F_{13}T_{13} + F_{14}T_{14}}{F_{13} + F_{14}} \tag{B.37}$$

$$T_{26} = \frac{F_{23}T_{23} + F_{25}T_{25}}{F_{23} + F_{25}} \tag{B.38}$$

$$T_{39} = \frac{F_{33}T_{33} + F_{36}T_{36} + F_{38}T_{38}}{F_{33} + F_{36} + F_{38}} \tag{B.39}$$

$$(UA)_A = \frac{1}{\frac{1}{11.7F_{12}^{0.8}} + \frac{1}{24.2F_1^{0.8}}} \tag{B.40}$$

$$(UA)_B = \frac{1}{\frac{1}{12.2F_{22}^{0.8}} + \frac{1}{21.4F_3^{0.8}}} \tag{B.41}$$

$$(UA)_C = \frac{1}{\frac{1}{12.0F_{24}^{0.8}} + \frac{1}{23.0F_7^{0.8}}} \tag{B.42}$$

$$(UA)_D = \frac{1}{\frac{1}{13.0F_{32}^{0.8}} + \frac{1}{25.0F_2^{0.8}}} \tag{B.43}$$

$$(UA)_E = \frac{1}{\frac{1}{10.9F_{35}^{0.8}} + \frac{1}{23.5F_5^{0.8}}} \tag{B.44}$$

$$(UA)_F = \frac{1}{\frac{1}{12.0F_{37}^{0.8}} + \frac{1}{24.0F_8^{0.8}}} \tag{B.45}$$

B.4 Simulation Results

Since the heat-exchanger network has four manipulated variables, it is not possible to present the simulation results as contour plots. Therefore, data from a simulation run of the closed loop system is presented as trend lines. Figure B.1 shows the setpoint changes for a simulation where all the manipulated variables, i.e., the split ratio of the valves, are initially at 0.3. The manipulated variables are changed by the RTO system to move the process closer to the optimum operating condition. After thirteen RTO steps the manipulated variables become stationary at their model optimum. The small changes after the thirteenth RTO steps are due to numerical and process noise added to the system measurement.

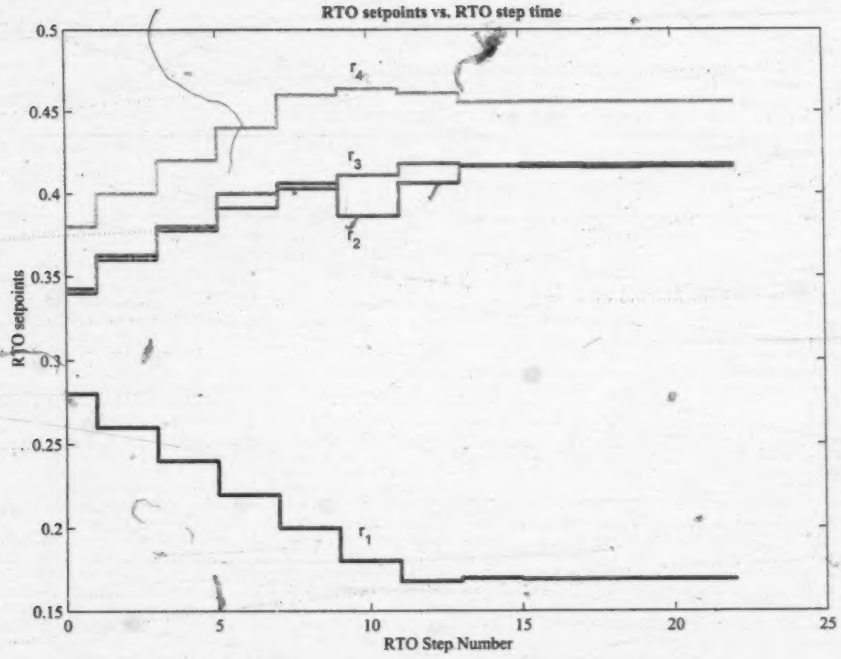

Figure B.1: RTO setpoints implemented by the optimizer

In every RTO step, the optimizer is trying to minimize the objective function (Equation B.8). Figure B.2 shows the gradual decline of the magnitude of the objective function. For the first few RTO steps, the operating point is significantly different from the optimum operating condition and therefore, the RTO system shows significant performance improvement. As the process moves closer to its optimum operating condition, the efficiency of the RTO system starts to decline.

Figure B.3 shows the plant-model mismatch angle calculated at the RTO intervals. The plot shows how the performance of the RTO system declines as it moves closer to the plant optimum. After the eighth RTO step, the angle between the gradients become larger than 90° indicating that the RTO system loses process performance in all of the RTO steps following it.

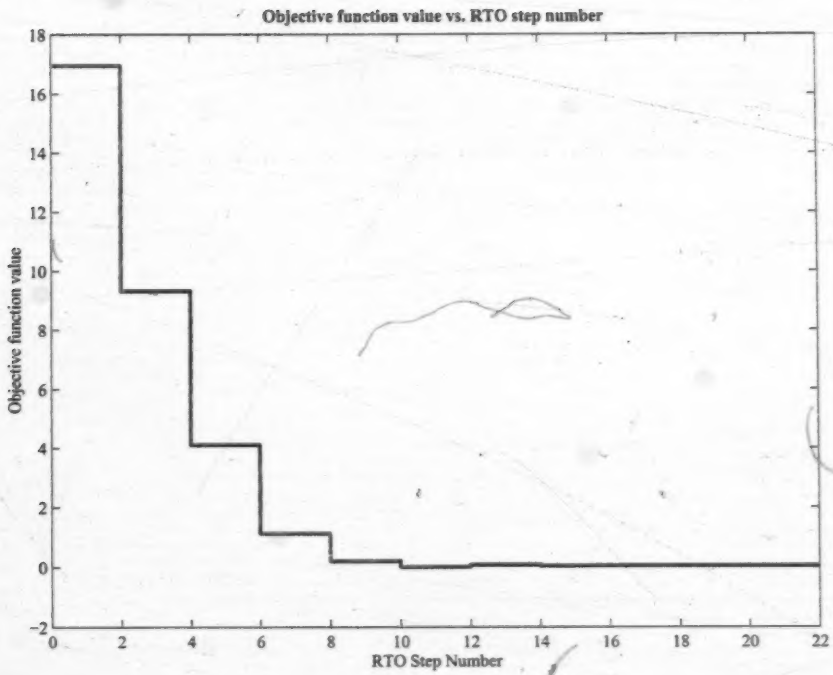

Figure B.2: Value of the objective function at different RTO steps

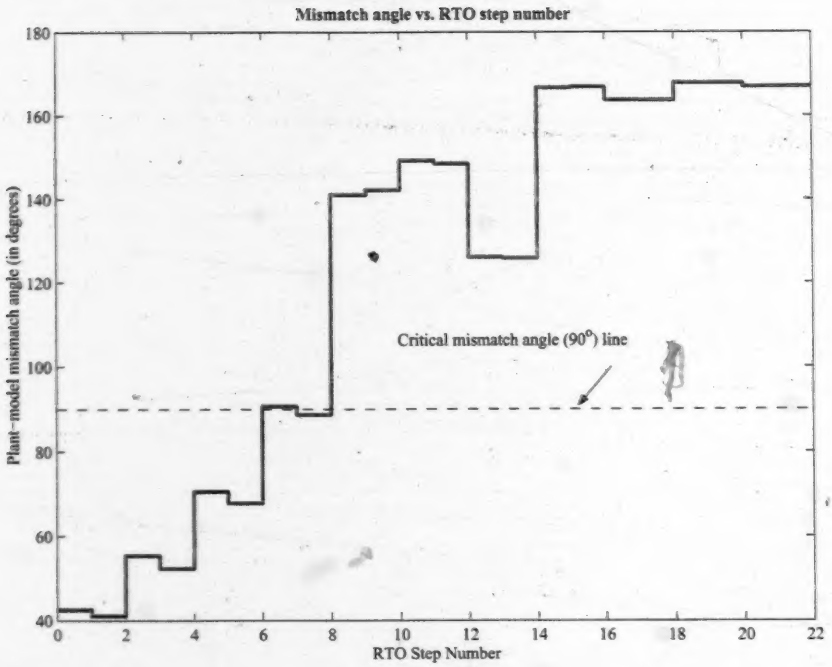

Figure B.3: Plant-model mismatch angle plotted against RTO steps